The Dissertation Committee for Chia-I Lin
certifies that this is the approved version of the following dissertation:

Biosynthetic Studies of Lincomycin A, a Thiosugar-Containing Natural Product

Committee:

Hung-wen Liu, Supervisor

Eric V. Anslyn

Guangbin Dong

Walter L. Fast

Christian P. Whitman

Adrian T. Keatinge-Clay

Biosynthetic Studies of Lincomycin A, a Thiosugar-Containing Natural Product

by

Chia-I Lin, B.S., M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2015

Dedicated to the inspirational teacher, Wu, Ching-Sung.

Acknowledgments

I would like to express my deepest appreciation and thanks to my advisor, Dr. Hung-wen (Ben) Liu, for being a tremendous mentor for me. I would like to thank Ben for encouraging my research and for allowing me to grow as a research scientist. I would also like to thank my committee members, Professor Eric Anslyn, Professor Guangbin Dong, Professor Walter Fast, Professor Christian Whitman, and Professor Adrian Keatinge-Clay for letting my defense be an enjoyable moment and for all the inspiring comments and brilliant suggestions. A special thank you to Dr. Yung-Nan Liu, who helps to purify some of the proteins in my projects and often provides good foods with amazingly healing effect.

I am grateful to the former Liu's lab members, including postdocs, Steve, Yasushi, Hui, Jordi, Reid and Chih-Hau, and graduate students, Wei-Chen, Sei Hyun, Hak Joong, Namho, Jake, Eta, Mel, Eita, Grace, Meilan, Chris, Jess and Ying. I am sincerely thankful to Eita for introducing me to the thiosugar project. Sei has been a tremendous help for me in both course work and organic synthesis. I would like to thank Mel, Grace and Meilan for being my friends during these years. Finally, I would like to express my gratitude to Reid and cherish the time we worked together.

I would like to thank current lab members as well. They are postdoc, Mark, and graduate students, Anthony, Ke-Yi, Hak Joong, Byung-sun, Yeonjin, Geng-Min, Jung-Kuei, Aoshu, Shao-An, Jon, Richiro and Yu-Cheng. I thank postdoc Mark for helping me learning LaTeX and his spot on suggestions for my posters and presentations. Anthony always shared with me fruitful information about novels, movies and research techniques. Ke-Yi has been a good friend and an in-

formative source of almost everything. Jon helped me to prepare some proteins for my project during his rotation period. Aoshu and Shao-An provided the synthetic compounds which are important for characterization of enzyme functions. Richiro joined the lincomycin A project and I am really grateful for that. Yeonjin is my dearest neighbor in the office.

I am using this opportunity to express my gratitude to the staff of the facilities in the University of Texas at Austin, including Dr. Ben Shoulders, Steven Sorey and Angela Spangenberg in the Nuclear Magnetic Resonance Lab and Dr. Karin Keller, Dr. Ian Riddington and Jordan Dinser in the Mass Spectrometry Facility of the Department of Chemistry, Dr. Stony H. Lo in the Analytical Instrumentation Facility of the College of Pharmacy, and Dr. Maria D. Person and Andre Bui in the Proteomics Facility of the Institute for Cellular and Molecular Biology.

Finally, I want to thank my parents and younger sister in Taiwan. Without their support and understanding, I would not be able to accomplish these works. I would also like to thank my boyfriend Chao-Yeh for helping me keep a sense of humor and happiness when I struggled. This dissertation is dedicated to Mr. Wu, Ching-Sung who was my mentor in junior high school, as he inspired me to explore the beauty of science.

Biosynthetic Studies of Lincomycin A, a Thiosugar-Containing Natural Product

Publication No. _____

Chia-I Lin, Ph.D.
The University of Texas at Austin, 2015

Supervisor: Hung-wen Liu

The dissertation describes biosynthetic studies of lincomycin A, a thiosugar-containing natural product. Lincomycin A was isolated from *Streptomyces lincolnensis* and has been used clinically as an effective antibiotic for more than 30 years. The structure of lincomycin A contains an *N*-methyl-4-propyl-L-proline moiety and a thiooctose. Although the clinical importance of lincomycin A has led to extensive efforts to optimize the fermentation conditions for its production, how it is biosynthesized, especially the construction of the thiooctose, remains poorly understood. The goal of my thesis research aims to identify the biosynthetic precursors of lincomycin and investigate the mechanism of the transformations involved in the biosynthetic pathway.

Based on the results of early feeding experiments and bio-informatic analysis of the biosynthetic gene cluster, a biosynthetic pathway of lincomycin could be proposed. The enzymes predicted to be involved in the biosynthetic pathway were purified in order to examine their functions. Organic synthesis was utilized to construct the proposed sugar substrates and the anticipated products to support the activity and the structural confirmation of the enzymatic products. With this approach, ribose and fructose were found to be the precursors that are assembled

in a transaldol condensation reaction catalyzed by LmbR to yield the eight-carbon sugar backbone of lincomycin. The stereochemistry of the enzymatic products was confirmed by the comparison with various synthetic stereoisomers.

The later parts of this research focus on the subsequent modifications of the octose precursor. Our studies led to the identification of a key intermediate, guanosine diphosphate-activated octose, which was formed through a kinase-phosphatase cascade reaction, catalyzed by LmbP (kinase), LmbK (phosphatase) and LmbO (guanylyltransferase). Further transformations of this GDP-sugar intermediate including epimerization, dehydration and transamination were also demonstrated via the catalysis of enzymes LmbM, LmbL/Z and LmbS, respectively.

How the C-1 sulfur is incorporated into the highly decorated GDP-octose is the most intriguing question in the biosynthesis of lincomycin A. Sequence analysis of the lincomycin gene cluster showed the presence of a putative *S*-glycosyltransferase, which likely catalyzes the displacement of GDP with an appropriate sulfur donor. This proposal has been verified recently by others. In addition, the amide bond formation between the proline moiety and thiooctose was shown to be catalyzed by a protein complex consisted of LmbC, LmbN and LmbD. This process resembles the biosynthesis of nonribosomal peptides.

In conclusion, this work unravels the biosynthetic origins of lincomycin A and also characterizes the full sequences of the enzymatic transformations of the octose intermediate. The significance of such biosynthetic studies of glycosylated natural products like lincosamide derivatives is due to the fact that these natural products are bioactive and pharmaceutically useful. Knowing their biosynthetic pathways gives us the opportunity to manipulate the biosynthetic machineries. Modifying the structure of the sugar components holds promise for varying their

biological activities. Moreover, the information about the biosynthetic precursors and the intermediates is useful for the optimization of the fermentation condition.

Table of Contents

List of Tables

List of Figures

Chapter 1

Background and Significance*

Carbohydrates are essential biomolecules for all types of living organisms. A wide variety of sugar moieties can be found in lipopolysaccharides, glycoproteins, and numerous secondary metabolites. In many cases, the glyco-component is an important determinant of bioactivity for the parent compound.[1,2] Given that glycosylation is a recurrent feature in natural products, one could reason that the diverse biological activities and molecular targets of glycosylated compounds may derive from the the great structural complexity of those sugar components.

The diverse structures of monosacchrides are largely resulted from functional group transformations, including oxidoreduction, isomerization, methylation, acetylation and transamination, before or after glycosyl transfer. The complexity observed in sugar structures is further enhanced by the heteroatom incorporation and high-carbon chain attachment. While carbohydrate is originally named for its common chemical composition, $(CH_2O)_n$, numerous sugars are found to contain covalent bonds between carbon and heteroatoms, such as nitrogen, sulfur, halogens and arsenic. For example, several sugars derivatives with antibiotic activities are found to contain C–N bonds in the forms of amine or oxidized amine.[3,4] A few naturally occurring thiosugars, in which the C–S bonds are present, have also been identified as pharmaceutically useful secondary metabo-

*Partial content of this chapter was published as "The biosynthesis of nitrogen-, sulfur-, and high-carbon chain-containing sugars" Lin, C.-I; McCarty, R. M.; Liu, H.-w., *Chem. Soc. Rev.*, **2013**, *42*, 4377–4407. This review was written by C.-I. L. and R. M. M., who contributed equally, and edited by H.-w. Liu.

lites.[5,6] Moreover, the carbohydrate intermediates containing C–X (X = F or Cl) bonds were discovered during the characterization of enzymes capable of catalyzing halogenation reactions.[7] Arsenosugars with C–As bond represent one of the major arsenicals in seafood.[8] Interestingly, although sugar phosphates are critical components in metabolism and phosphonic/phosphinic natural products are widely distributed,[9] there is no C–P bond containing sugar discovered so far. The focus of this chapter is the biosynthesis of sulfur-containing sugars. This class of unusual sugars are interesting due to the existence of the thio-substituents which pose impact on the molecular properties including the elctronegativity, hydrogen bonding ability and steric hindrance. In addition to the thio-constituents, some sugars feature "high-carbon" chains as their backbones. A greater than usual carbon number ($>$ 6) not only extend the length of these unusual sugars, but also allows multiple site functionalization and fused ring formation.

As far as glycosylated natural products are concerned, research efforts in the biosynthesis of deoxysugars and the mechanism of glycosyltransferases have made it possible for manipulation of natural products glycodiversification.[10,11] The construction of the sulfur- and high-carbon chain-containing sugars are less understood due to their relatively rare natural occurrences. Thus, exploring the biosynthesis of these natural products may aid in the identification of novel compounds with improved pharmacological properties. In this chapter, the known sulfur- and high-carbon chain-containing sugars will be introduced and current understanding regarding their biosynthesis will also be discussed.

1.1 Sulfur-containing Sugars

Sulfur is the fifth most abundant element in living organisms, after oxygen, carbon, hydrogen and nitrogen. Sulfur-containing molecules are widely distributed in natural products, including amino acids, cofactors, antioxidants, nu-

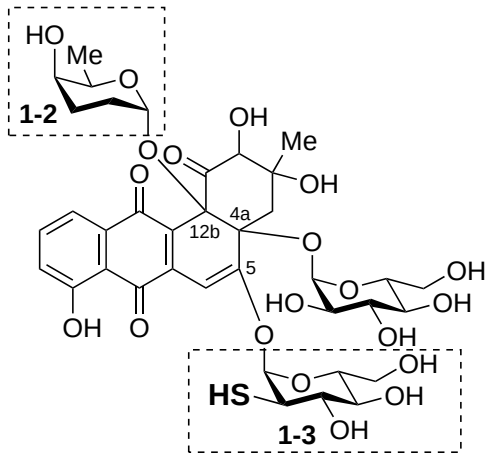
cleotides and numerous secondary metabolites.[5] However, only a few naturally occurring thiosugars have been identified and the relative scarcity has hindered investigation regarding their biosynthesis and biological activities.[6,12] The term, thiosugar, is used here to describe a carbohydrate molecule that carries a sulfur-containing functional group or is connected to other molecules via a sulfide linkage. In this section, the known thiosugars are categorized into four classes based on the oxidation state of the sulfur-containing functional groups: thiols, sulfides, sulfonium ions and sulfonic acids. Their biosynthetic pathways and the enzymes catalyzing the sulfur incorporation step in each pathway are discussed based on the data that are currently available.

1.1.1 Group I: Thiosugars Carrying a Free Thiol Group

The only two members of group I are the angucycline-type antibiotics rhodonocardin A (**1-1**) and B isolated from *Nocardia* sp. No. 53 (Figure 1.1).[13] In rhodonocardin A, the benz[*a*]anthraquinone skeleton is decorated with three sugar units, rhodinose (**1-2**), 2-thioglucose (**1-3**), and glucose, which are joined via *O*-glycosidic bonds to C-12b, C-5, and C-4a, respectively. Rhodonocardin B contains only the first two sugars. The biosynthesis of rhodonocardins has not been studied.

1.1.2 Group II: Thiosugars Carrying a Sulfide Linkage

Sulfides are compounds containing C–S–C connectivity. Compared to the free thiols, which display a much higher tendency toward oxidation and stronger nucleophilicity, sulfides are more resistant to chemical reactions and are generally more stable. For this reason, most of the known naturally occurring thiosugars belong to group II. In this class of compounds, the sulfur atom either exists in the endocyclic ring, or serves as a linkage between the sugar molecule and the aglycone. Compounds within group II are further divided into five subgroups



Rhodonocardin A (**1-1**)

Figure 1.1: The angucycline-type antibiotic rhodonocardin A.

according to their overall chemical structures.

1.1.2.1 Thiocycle Ring

The first example of a thiocyclic sugar, 5-thio-D-mannose (**1-3**), was isolated from the marine sponge *Clathria pyramida* in 1987 (Figure 1.2). This is so far the only free thiocycle-containing monosaccharide identified in nature.[14] It has been speculated that 5-thio-D-mannose may be produced by a sulfate reducing bacterial symbiont of *C. pyramida*, rather than by *C. pyramida* itself.

The albomycin δ_1 (**1-4**), isolated from *Streptomyces subtropicus*, contain a 6-amino-6-deoxy-4-thio-heptofuranose uronic acid nucleoside (**1-5**) and a ferrichrome siderophore moiety (**1-6**).[15,16] These two subunits are linked via a serine residue (Figure 1.2). The iron-chelating siderophore can be recognized by the bacterial siderophore-dependent iron acquisition system. This allows albomycins to be actively transported across the membranes leading to their designation as "Trojan horse" antibiotics.[17] The seryl-**1-5** has been demonstrated to be a potent inhibitor against bacterial seryl-tRNA synthetase (SerRS).[18] The presence of sulfur in albomycin is critical for the observed antimicrobial activity, because the oxygen-ring

analogue of albomycin is totally inactive.[19]

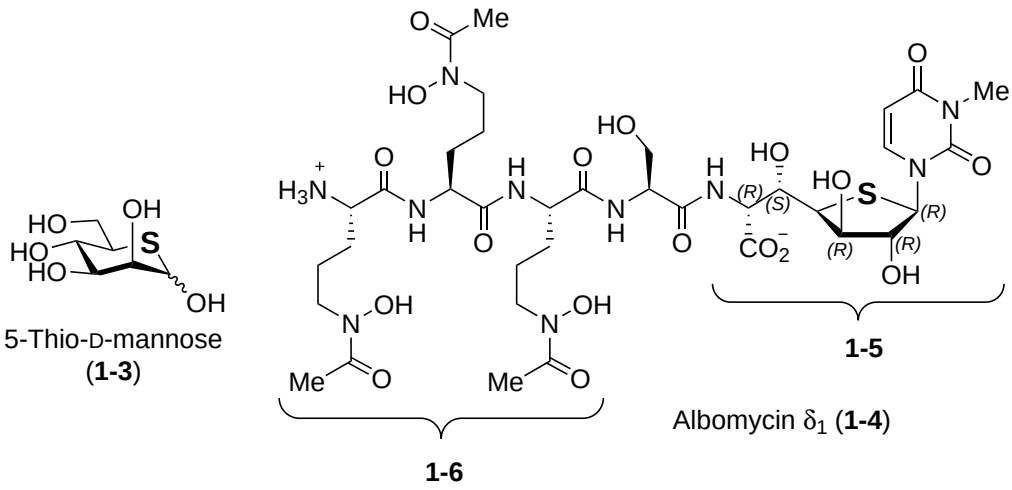


Figure 1.2: Thiosugars containing a thiocycle ring.

The identification of the albomycin biosynthetic gene cluster was recently reported by Chen and coworkers.[20] It has been speculated that the sulfur incorporation step proceeds via a mechanism analogous to that employed by BioB, the biotin synthase (BioB) (Figure 1.3A).[21,22] BioB is a member of the radical SAM enzyme superfamily.[23] Radical SAM enzymes harbor a [4Fe–4S] cluster, which after reduction can trigger the reductive cleavage of SAM to give methionine and the 5′-deoxyadenosyl radical (**1-7**). The latter is a powerful oxidant that can abstract a hydrogen atom from substrates to yield radical intermediates, which then undergo radical-mediated transformations. In BioB, the 5′-deoxyadenosyl radical (**1-7**) abstracts hydrogen from dethiobiotin (**1-8**) and the resulting substrate radical (**1-9**) takes sulfur out of a [2Fe–2S] cluster bound in the active site of BioB. Since *abmM* appears to encode a radical SAM enzyme with both the [4Fe–4S] and [2Fe–2S] cluster bound, the same sulfur insertion mechanism may be operative in AbmM (**1-10**→**1-11**, Figure 1.3B). It should be noted that the biosynthesis of 4-thioheptofuranose uronic acid moiety **1-5** will be discussed in Section 1.2.1.

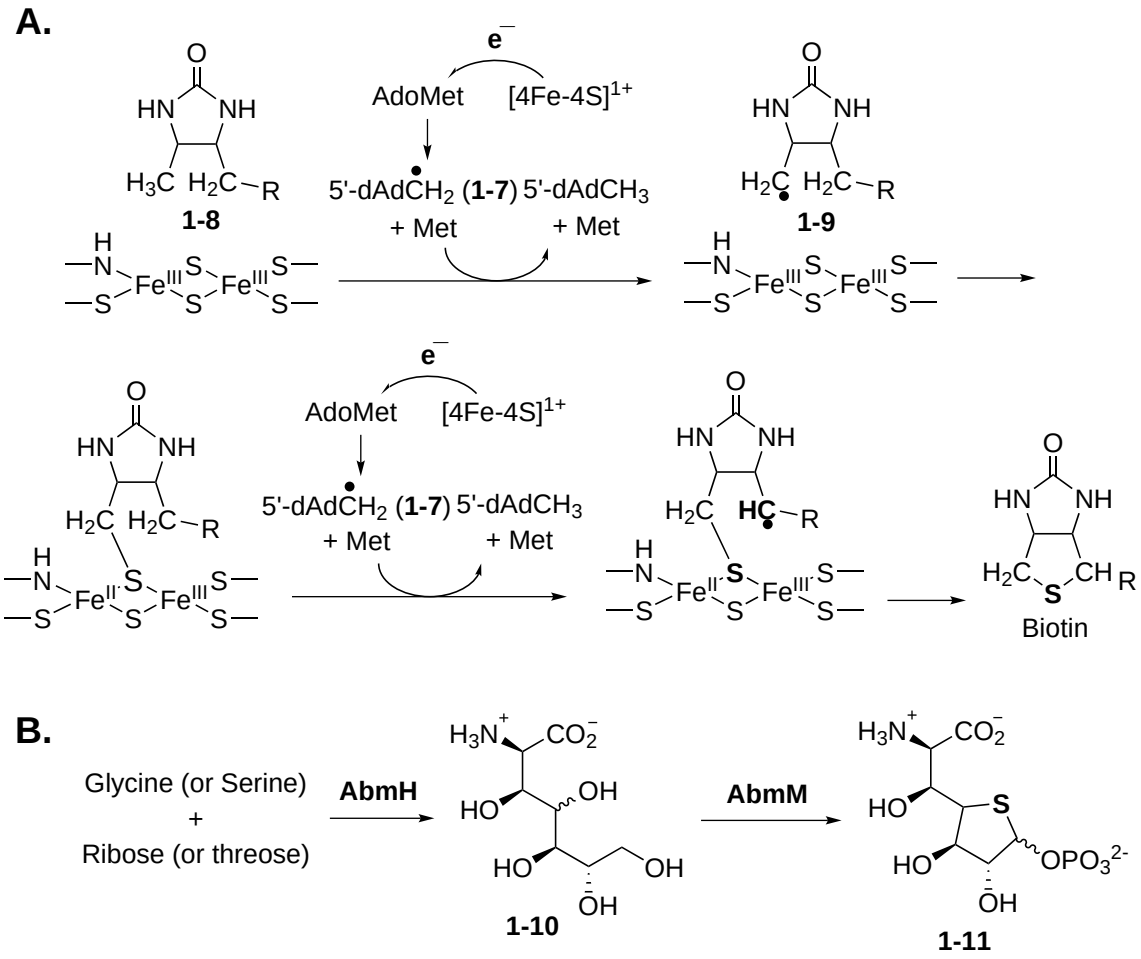


Figure 1.3: (A) Proposed mechanism for BioB-catalyzed sulfur incorporation in biotin biosynthesis. (B) Proposed mechanism for sulfur incorporation in albomycin biosynthesis.

1.1.2.2 Angucycline-type Antibiotics

BE-7585A (**1-12**) from *Amycolatopsis orientalis* subsp. *Vinearia* is a constitutional isomer of the aforementioned group I thiosugar-containing rhodonocardin A (**1-1**).[24] In BE-7585A, 2-thioglucose (**1-3**) and glucose are joined together as a 1,1′-*O*-linked disaccharide that is attached to the C-5 of aglycone via a thioether linkage. The biosynthetic studies of BE-7585A led to the identification and characterization of a 2-thioglucose-6-phosphate synthase, BexX.[25] This enzyme shows significant sequence homology to the thiazole synthase ThiG, which catalyzes the C–S bond formation of the thiazole moiety (**1-13**) during thiamine biosynthesis (Figure 1.4).[26] It is therefore proposed that BexX may perform an analogous function in the biosynthesis of 2-thioglucose (Figure 1.5). Using BexX and glucose-6-phosphate (G6P, **1-14**), a covalent adduct (**1-15**) between Lys-110 of BexX and C-1 of a 2-ketosugar derived from G6P was detected in the crystal structure.[27] This result strongly supports that BexX is the enzyme responsible for priming G6P to accept a nucleophilic sulfur donor at C-2 (**1-15**→**1-16**, Figure 1.5). However, no potential sulfur delivery protein, such as ThiS, which works along with ThiG in thiazole biosynthesis (Figure 1.4), is encoded in the *bex* gene cluster. Through genome sequencing, it was later discovered that BexX actually co-opts the sulfur-delivery systems from primary metabolism, such as those used in cysteine and molybdopterin biosynthesis to deliver the sulfur atom to the sugar substrate bound in the active site of BexX.[27]

1.1.2.3 Lincosamides

Lincomycin A (**1-17**), isolated from *Streptomyces lincolnensis*, displays antibiotic activity against Gram-positive bacteria.[28,29] It has been suggested that lincomycin A blocks protein synthesis by binding to the peptidyltransferase domain of the 50S ribosomal subunit.[30] Lincomycin A contains a 1-thiooctose core (**1-20**),

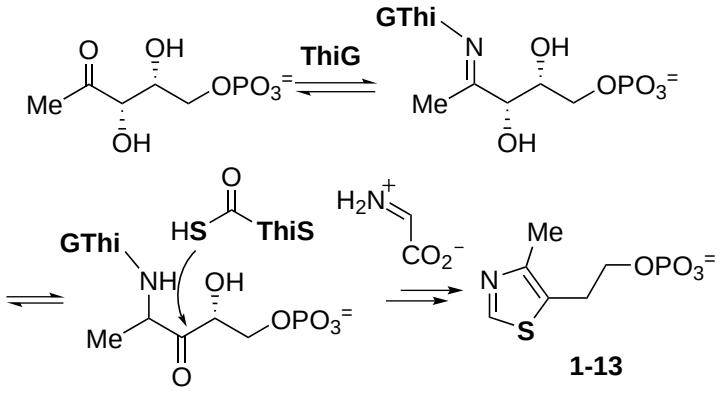


Figure 1.4: ThiG-catalyzed thiazole formation.

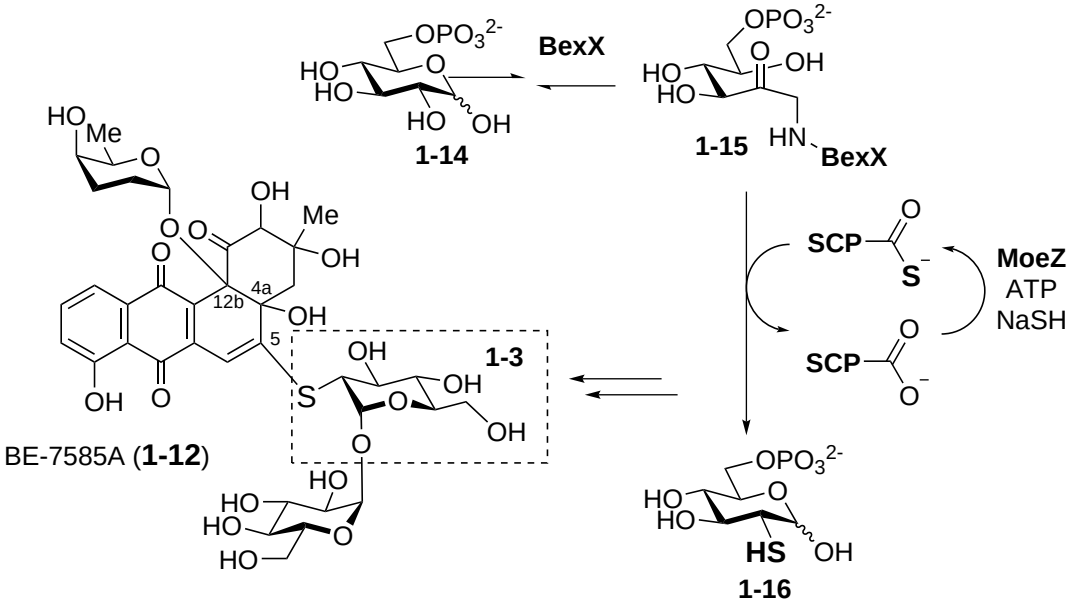


Figure 1.5: Biosynthesis of 2-thioglucose in BE-7585A

which is also found in Bu-2545 (**1-18**) from *Streptomyces* strain H230-570 and celesticetin (**1-19**) from *Streptomyces caelestis* (Figure 1.6).[31–33] This eight-carbon skeleton of lincomycin makes it also recognized as a high-carbon sugar (Section 1.2.2). Progress has been made in learning the biosynthesis of lincomycin and celesticetin. The available data will be discussed in later chapters of this dissertation.

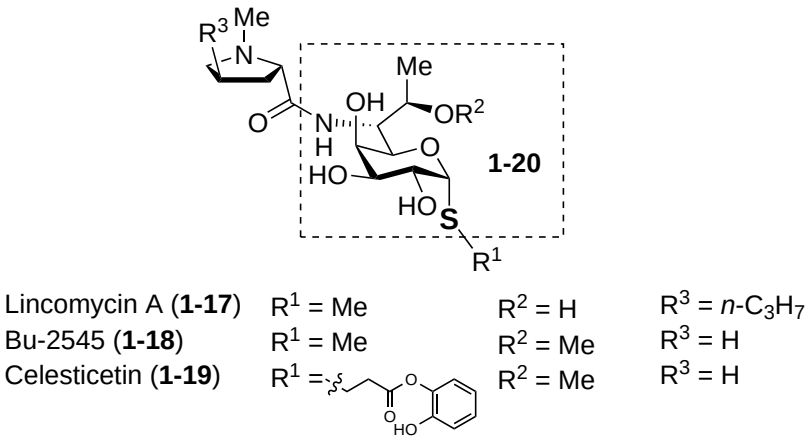


Figure 1.6: The family of lincosamide-containing antibiotics.

1.1.2.4 Endiyne-type Natural Products

Several 4-thiosugars (**sulfide-xx**) have been found as structural components of endiyne-containing natural products: calicheamicin (**1-21**) from *Micromonospora echinospora* ssp. *Calichensis*,[34] namenamicin from *Polysyncraton lithostrotum*,[35] esperamicin from *Actinomadura verrucosospora*,[36] and shishijimicin from *Didemnum proliferum*.[37] Based on the analysis of the calicheamicin biosynthetic gene cluster, the translated product of the *calS4* (or *calK*) gene, which encodes a putative C–S lyase, has been proposed to catalyze the conversion of a TDP-2,6-dideoxy-4-ketosugar (**1-22**) to the corresponding 4-thiosugar (**1-23**, Figure 1.7).[38] However, this hypothesis has not been verified and it is unclear whether other proteins are also required for the sulfur incorporation step.

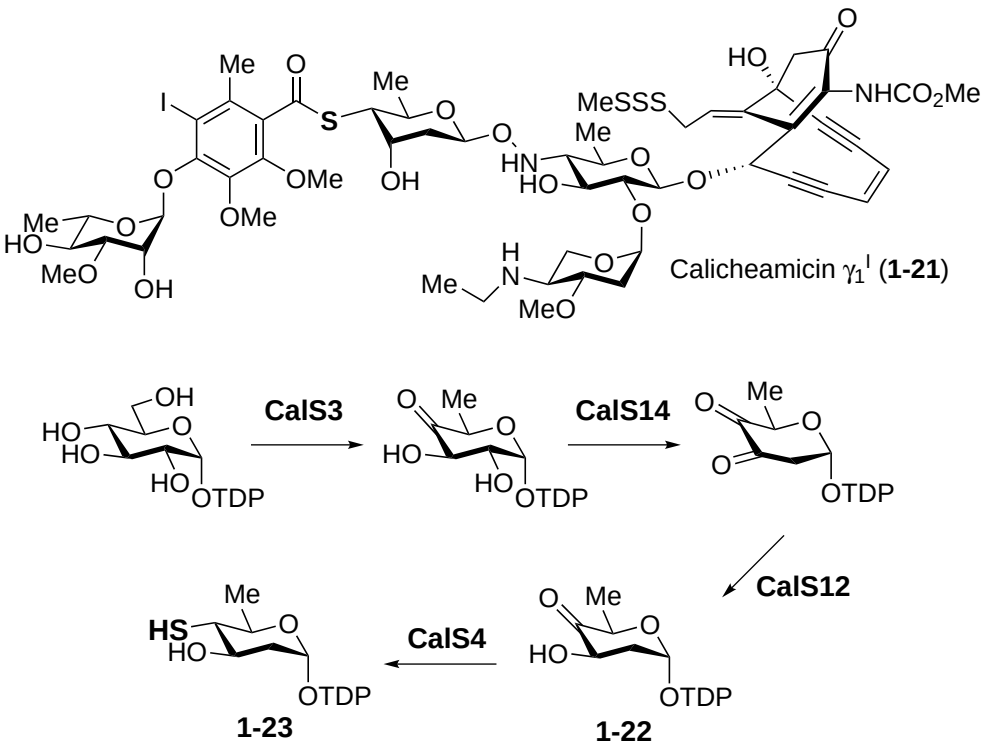


Figure 1.7: Proposed biosynthetic pathway of 4-thiosugar in calicheamicin.

1.1.2.5 Glucosinolates

Glucosinolates (**1-24**) are 1-thiosugar-containing natural products produced by *Brassicales* plants and are utilized for defense against insects and pathogens.[39] Damage to the plant tissues will trigger the hydrolysis of the C–S glycosidic bond of glucosinolates by myrosinases to release unstable thiohydroximate-*O*-sulfates. The sulfate moiety is then lost via a spontaneous Lossen rearrangement resulting in the formation of various toxic compounds including isothiocyanates. Structurally, glucosinolates are β-thioglucoside *N*-hydroxysulfates containing a side chain that is derived from an amino acid.

The biosynthesis of glucosinolates has been extensively investigated and reviewed.[40–42] As depicted in Figure 1.8, an amino acid precursor is converted to a nitrile oxide (**1-25**) or aci-nitro (**1-26**) intermediate via consecutive oxidations catalyzed by cytochrome P450 monooxygenases. The critical sulfur incorpora-

tion step has been proposed to occur nonenzymatically with cysteine as the sulfur donor. Subsequent cleavage of the thio-conjugate (**1-27**) by PLP-dependent C–S lyase SUR1 yields thiohydroximate (**1-28**), which is then glycosylated and sulfonated to yield the glucosinolate core. Although cysteine has long been considered to be the sulfur source in glucosinolate biosynthesis, some reports have indicated that glutathione (GSH) might be a more plausible sulfur donor and a specific GSH-*S*-transferase (GST) may catalyze the sulfur incorporation step (Figure 1.8).[43–45]

1.1.3 Group III: Thiosugars Carrying a Sulfonium Ion

Isolated from the medicinal herb *Salacia reticulate*, salacinol (**1-29**) and kotalanol (**1-30**) are both naturally occurring thiosugars that contain a sulfonium ion moiety (Figure 1.9).[46,47] Several analogues, including ponkolanol and salaprinol, were later found from other *Salacia* species.[48] The unique zwitterionic structure of salacinol (**1-29**) and its analogues includes a 1-deoxy-4-thio-D-arabinofuranosyl cationic core with 1-deoxy-aldosyl-3-sulfate as the anionic branched-chain. These compounds exhibit high α-glucosidase inhibitory activity and have consequently emerged as anti-diabetic drug candidates.[48] While extensive structure activity relationship studies have been carried out,[49] the biosynthetic pathways for the formation of these sulfonium-containing sugars remain unknown.

1.1.4 Group IV: Thiosugars Carrying a Sulfonic Acid

The unusual sulfosugar was found as a component of sulfoquinovosyl diacylglycerol (SQDG, **1-31**).[50] SQDG comprises about 47% of total lipids in chloroplast membranes of most photosynthetic organisms.[51] Feeding studies using [^{35}S]-sulfate traced the sulfolipid biosynthetic pathway to a nucleotide-activated sulfoquinovose intermediate (**1-32**, Figure 1.10).[52] It was later identified that *sqdB* en-

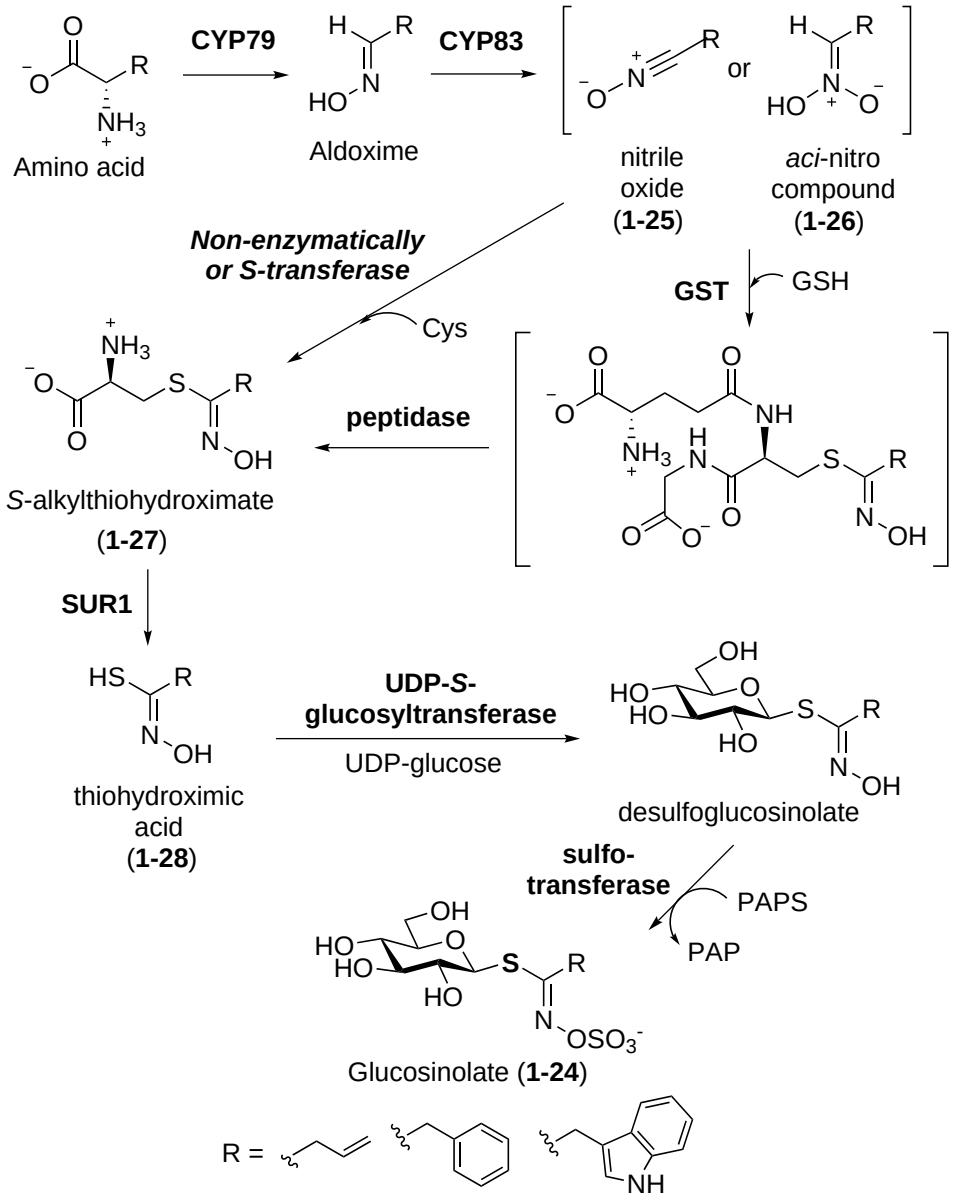


Figure 1.8: Proposed pathway for the biosynthesis of glucosinolate.

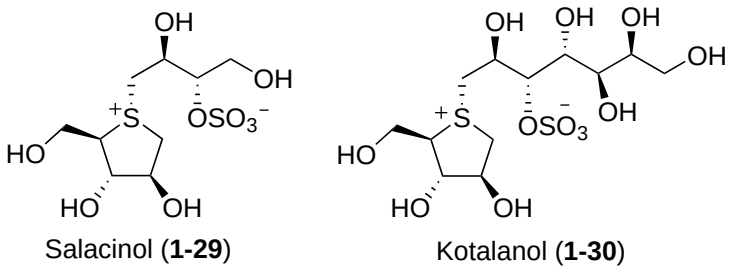


Figure 1.9: Thiosugars containing a sulfonium ion moiety.

codes the enzyme catalyzing the insertion of inorganic sulfite into UDP-glucose.[53–55] However, the exact physiological sulfite source and the detailed mechanism of sulfonyl incorporation awaits further investigation.

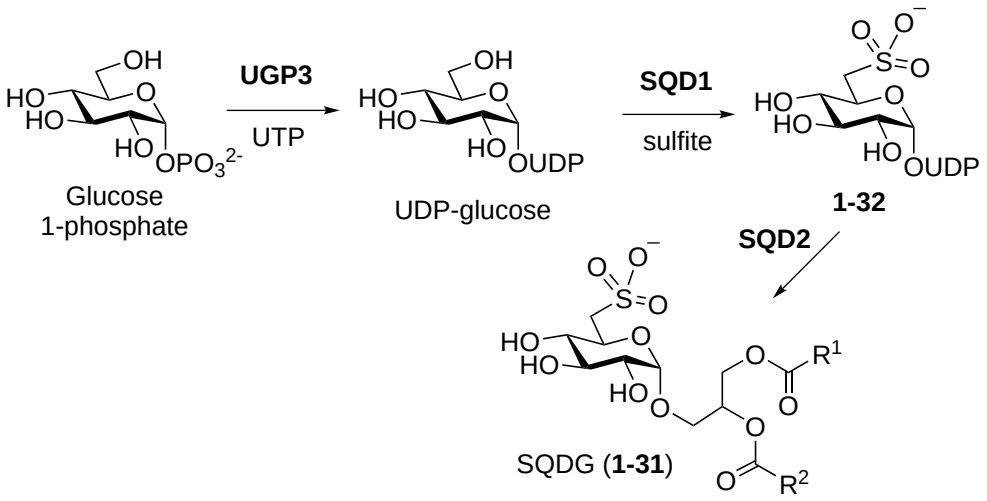


Figure 1.10: Proposed pathway for the biosynthesis of sulfoquinovosyl diacylglycerol.

1.2 High-carbon Chain-containing Sugars

Unusual sugars with high carbon numbers are found in an array of microbially produced antibiotics and in the outer walls of some Gram-negative bacteria. They are generally formed via condensation reactions between the C-5 and C-6 sugars of primary metabolism, or derivatives thereof, and an exogenous carbon source. The known mechanisms employed for condensation during the biosynthesis of high-carbon sugars are as varied as the mechanisms for biological C–C bond formation in general. It should be noted that highly oxygenated hydrocarbons may cyclize to form a multi-hydroxylated ring structure, which appears as high-carbon chain-containing sugars. These high-carbon sugars recorded in the literature are classified by the carbon numbers and their biosynthesis are discussed in the following sections.

1.2.1 Heptoses

A number of heptoses have been identified as essential components of the lipopolysaccharides (LPS), the amphipathic glycolipids found as a part of the outer membrane of Gram-negative bacteria.[56] The common structure of LPS comprises the lipid A with a bisphosphorylated and acylated β-(1$\rightarrow$6) glucosamine disaccharide, the inner-core part containing heptoses and 3-deoxy-D-*manno*-oct-2-ulosonic acid (KDO), and a polysaccharide chain. Most of the LPS heptoses have a *glycero*-D-*manno* configuration. Biosynthetic pathways from D-sedoheptulose 7-phosphate to these heptoses have been characterized[57,58] and will be discussed in later chapters due to their similarity to the biosynthesis of lincomycin.

Liposidomycin (**1-33**), caprazamycin (**1-34**), and A-90289 (**1-35**) share a seven-carbon 5′-*C*-glycyluridine core (**1-36**).[59] These uridyl lipopeptides inhibit the enzyme MraY, phopho-MurNAc-pentapeptide translocase, involved in the bacterial peptidoglycan biosynthetic pathway.[59,60] Gene clusters have recently been identified for the biosynthesis of liposidomycin in *Streptomyces* sp. SN1061M (*lpm*), caprazamycin in *Streptomyces* sp. MK730-62F2 (*cpz*), muraymycin in *Streptomyces* sp. NRRL 30471 (*mur*), and A-90289 in *Streptomyces* sp. SANK 60405 (*lip*).[61–64] Among the four compounds, the biosynthesis of A-90289 has been extensively characterized (Figure 1.11). The pathway is initiated by the oxidation of uridine-5′-monophosphate (UMP) to uridine 5′-aldehyde (**1-37**$\rightarrow$**1-38**). This reaction is catalyzed by a non-heme iron(II)-dependent α-ketoglutarate dioxygenase, LipL.[65] Subsequently, a putative serine hydroxymethyltransferase (SHMT) encoded by *lipK* was shown to catalyze the condensation between uridine 5′-aldehyde (**1-38**) and the PLP-bound glycine.[66] The homologs of LipL and LipK can be found in most of uridyl lipopeptide biosynthetic pathways since they are essential for the construction of the heptose unit.

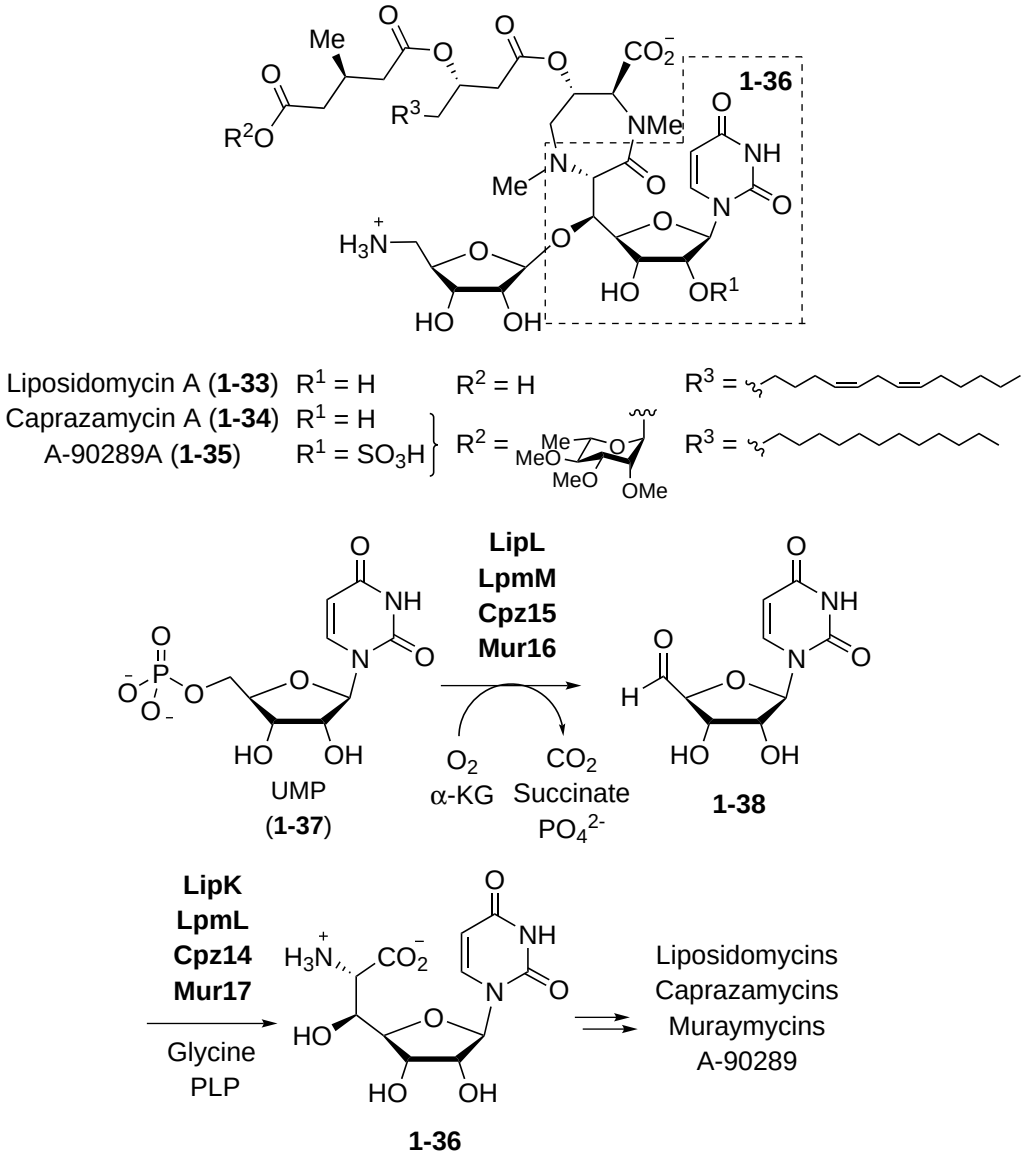


Figure 1.11: Biosynthetic pathway of 5′-*C*-glycyluridine moiety in liposidomycin and related compounds.

Albomycin δ_1 (**1-4**) has been briefly mentioned in Section 1.1.2.1. Although currently there is no biochemical data regarding the biosynthesis of the 4-thio-heptouronic acid moiety **1-5** in albomycin, three possible pathways have been proposed based on sequence analysis of the albomycin biosynthetic gene cluster.[20] As illustrated in Figure 1.12, in pathway A the reaction is initiated by the nucleophilic addition of cysteine sulfur to C-3 of phosphoenolpyruvate (PEP, **1-39**) to form a thioether adduct (**1-40**). This intermediate may undergo a series of enzymatic reaction to yield (**1-5**). An analogous reaction sequence is also proposed for pathway B where homocysteine instead of cysteine is used as the substrate. In the pathway C, either serine or glycine is joined with threose or ribose, respectively, via an intermolecular aldol condensation reaction. This is followed by sulfur insertion and cyclization to give **1-41**, which could then be phosphorylated at C-1 by the deoxynucleoside kinase homologue AbmG to give **1-5**.

The antitumor compound spicamycin (**1-42**) from *Streptomyces alanosinicus* 87-MT3 and its C-2 epimer septacidin (**1-43**) from *Streptomyces fimbriatus* were both isolated as mixtures of similar compounds carrying fatty acyl substituents of variable lengths.[67,68] The structures of both septacidin and spicamycin contain 4-amino-4-deoxyheptose (**1-45**) as does anicemycin (**1-44**) from *Streptomyces* sp. TP-A0648, which was identified recently.[69]

Amipurimycin (**1-46**), isolated from the culture media of *Streptomyces novoguineensis*, was shown to display antibiotic activity against *Pyricularia oryzae*, a fungal phytopathogen that causes sheath blight in rice plants.[70] The structure of amipurimycin includes an unusual branched heptose.[71] The similar heptose structure was also found in Miharamycin A (**1-47**), an adenine-containing nucleoside isolated from *Streptomyces miharaensis* SF-4890.[72] So far, there were no further information regarding the biosynthesis of these two antibiotics.

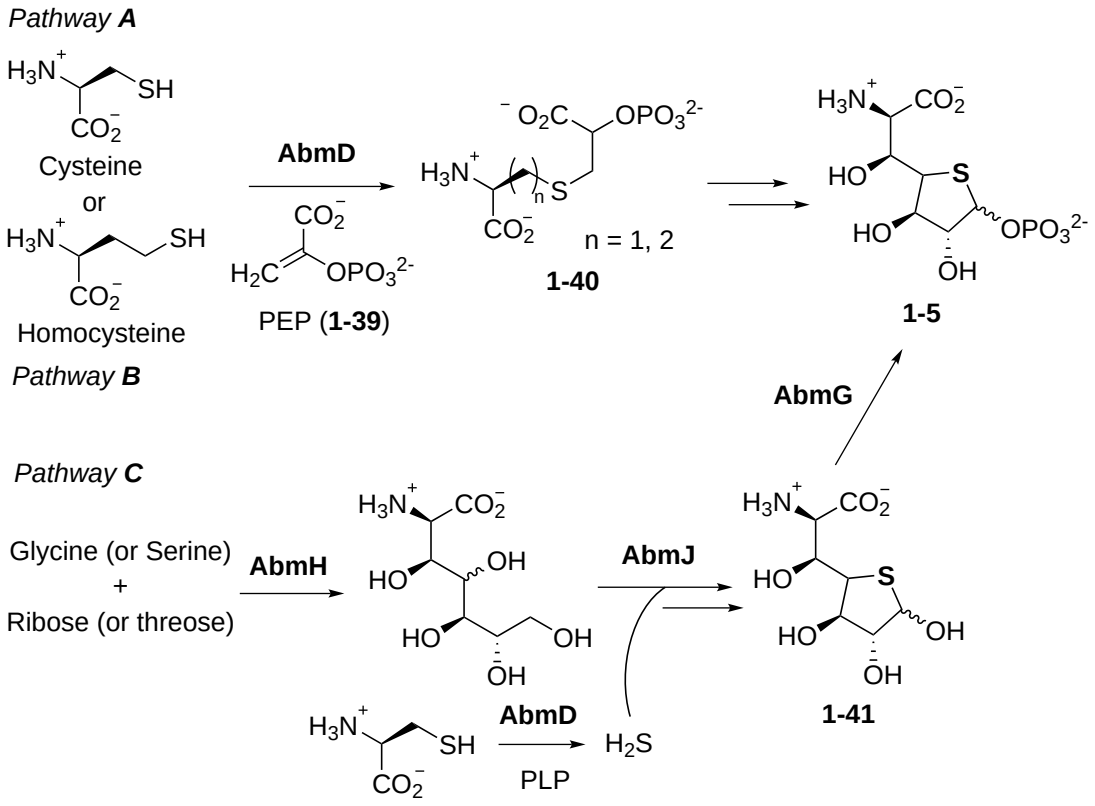


Figure 1.12: Proposed biosynthetic pathways of 4-thioheptouronic acid in albomycin.

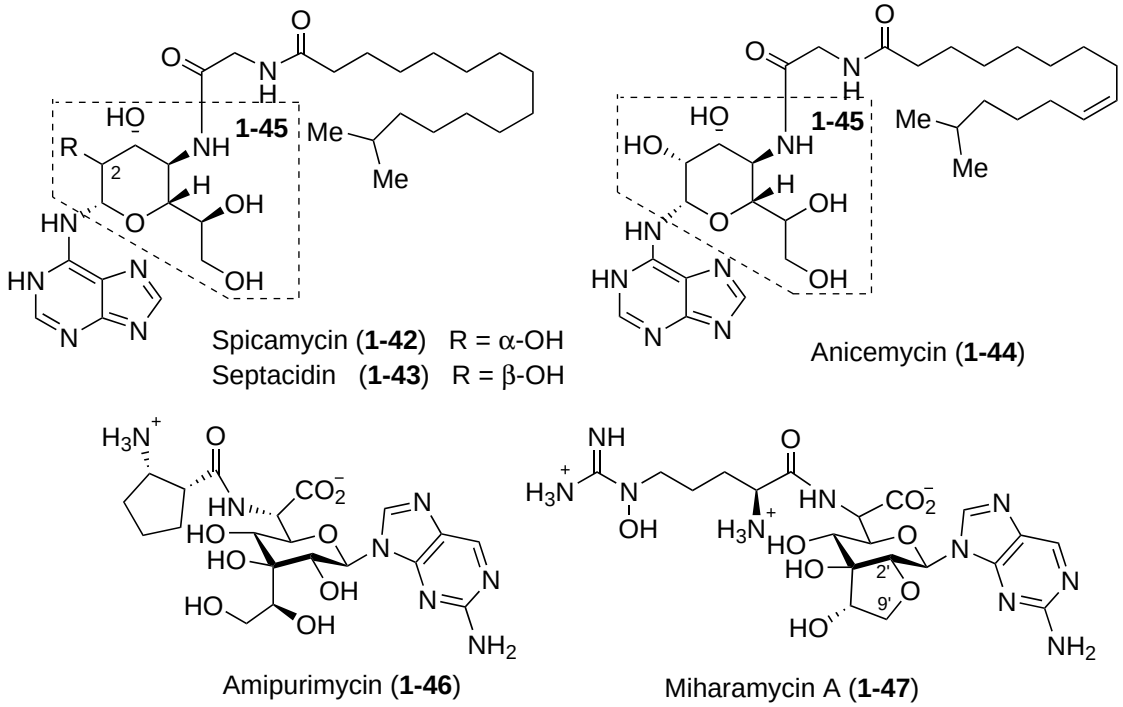


Figure 1.13: Other heptose-containing natural products.

1.2.2 Octoses

In addition to the bacterial heptoses described above (Section 1.2.1), KDO (**1-48**), is also an important component of the LPS in Gram-negative bacteria.[73] KDO is biosynthesized via an aldol condensation between arabinose-5-phosphate (**1-49**) and PEP (**1-39**) catalyzed by KDO-8-phosphate synthase (Figure 1.14).[74]

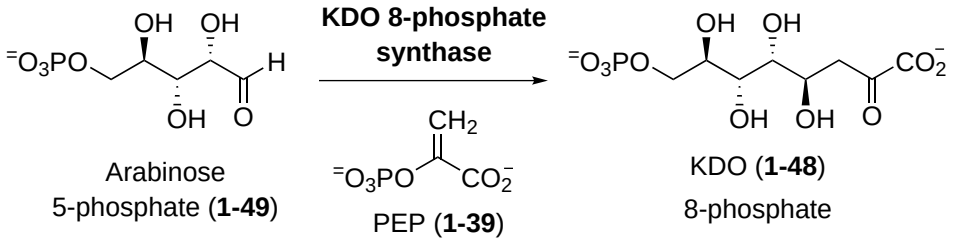


Figure 1.14: Biosynthetic pathway of KDO.

Octosyl acid A (**1-50**) is a bicyclic, anhydrooctose uronic acid nucleosides isolated from media of *Streptomyces cacaoi* var. *asoensis*.[75] This strain also produces polyoxin C (**1-51**), an active antifungal agent via inhibition of chitin synthase.[76] Feeding experiments using isotope labeled precursors revealed that octosyl acids and polyoxins may be derived from a common octose intermediate (**1-52**), which likely results from the condensation between PEP and a uridine derivative (Figure 1.15. Biosynthetic gene cluster involved in the production of polyoxins in *S. cacaoi* (*pol*) has been identified.[77] The gene product of *polA* was shown to catalyze the formation of 3′-enoylpyruvoyl-UMP from PEP and UMP. However, it remains unclear how the subsequent cyclization of **1-52** occurs and finally yields the bicyclic structure of octosyl acids.

Apramycin (**1-53**) is an aminoglycoside antibiotic that contains an unusual bicyclic octose sugar (**1-54**).[78] Labeling studies have indicated that the carbon atoms from C-1′ to C-6′ of the octose moiety are derived from glucose while C-7′ and C-8′ may originate from C-2 and C-3 of pyruvate.[79] Gene clusters for the biosyn-

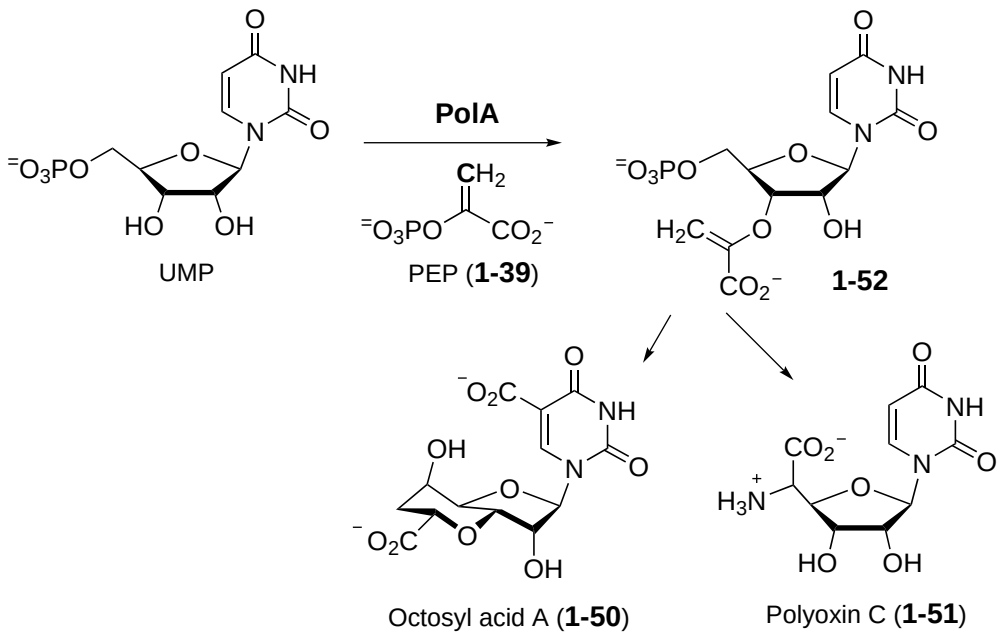


Figure 1.15: Proposed pathways for the biosynthesis of octosyl acid A and polyoxin C.

thesis of apramycin in *Streptoalloteichus tenebrarius* DSM 40477T and *Streptoalloteichus hindustanus* DSM 44523T have been identified (designated *apr*).[80] A possible biosynthetic pathway has been proposed based on the gene annotations in the *apr*. As shown in Figure 1.16, AprQ, a protein homolog of putative paromamine 6′-dehydrogenase, is predicted to oxidize the intermediate (**1-55**) to the corresponding aldehyde (**1-56**), which may undergo an aldol condensation with a two-carbon unit to give the octose sugar (**1-57**). However, the precise source of this two-carbon species and the mechanism of the octose formation remain unknown.

The lincosamide antibiotics were introduced earlier (Section 1.1.2.3). The biosynthetic precursors of the octose moiety in lincosamide have been shown to be D-fructose 6-phosphate (F6P) or D-sedoheptulose 7-phosphate as the three-carbon donor and D-ribose 5-phosphate (R5P) as the five-carbon acceptor. The detailed mechanism of octose formation will be further described in later chapters of this dessertation.

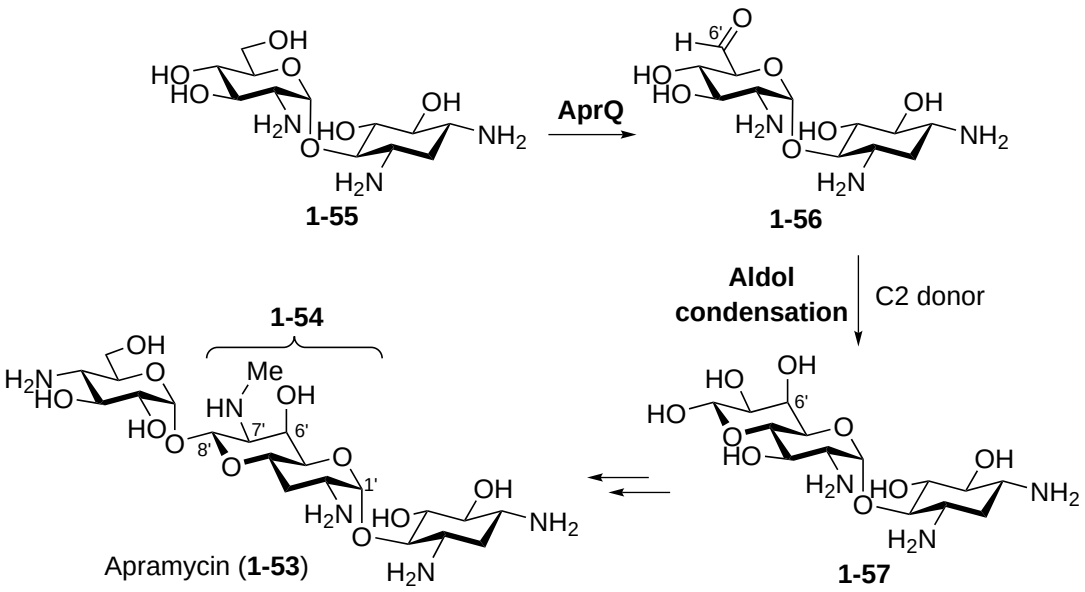


Figure 1.16: Proposed biosynthetic pathway of the bicyclic octose moiety in apramycin.

1.2.3 Nonose

As it has been mentioned in sections 1.2.1 and 1.2.2, there are several high-carbon sugars identified as crucial constituents of microbial cell walls. Perhaps the most widely recognized high-carbon sugar in this class is *N*-acetyl-neuraminic acid (NeuNAc, **1-58**). This nine-carbon sugar is a key component of the eukaryotic extracellular matrix, and is also found to a lesser extent in bacteria. In bacterial systems, it is formed via an aldol condensation between *N*-acetyl-mannosamine (ManNAc, **1-59**) and either PEP (**1-39**) or pyruvate by *N*-acetyl-neuraminic acid synthase or aldolase, respectively (Figure 1.17).[81,82] The reaction is very similar as the condensation catalyzed by KDO-8-phosphate synthase (Figure 1.14).

Griseolic acid A (**1-60**) is a cyclic AMP analog that contains an unusual nine-carbon bicyclic sugar, dicarboxylate glycone.[83] Griseolic acid A and its analogs were identified and isolated from culture media of *Streptomyces griseoaurantiacus* during a search for inhibitors of cyclic nucleotide phosphodiesterase. Feeding experiments showed that adenosine, oxaloacetate and succinate may be the biosyn-

thetic precursors for griseolic acid A, but no further characterization was performed to elucidate its biosynthetic pathway.[84]

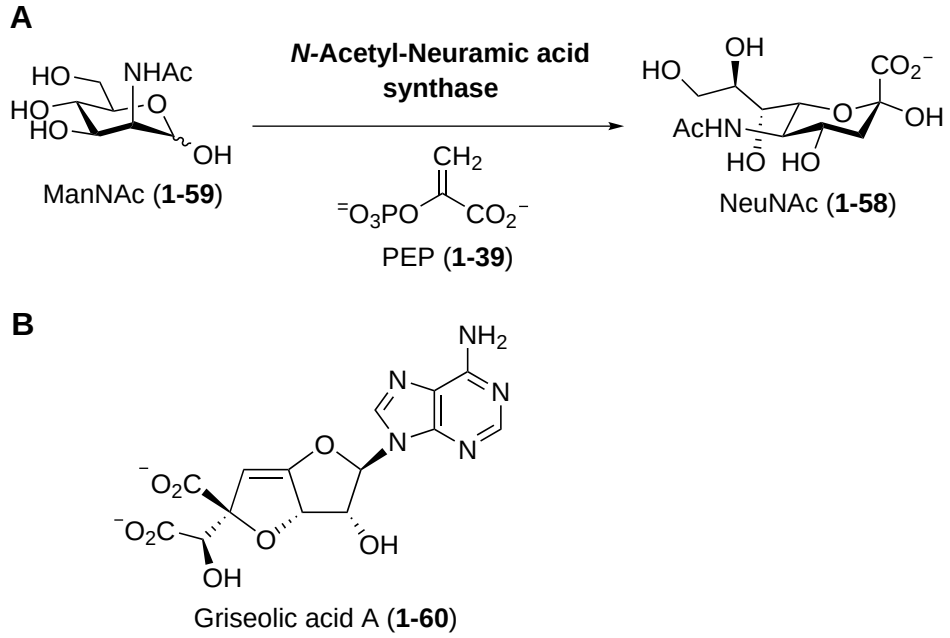


Figure 1.17: A. Biosynthetic pathway of *N*-acetyl-neuraminic acid. B. Griseolic acid A.

1.2.4 Decoses and Sugars Containing More Than Ten Carbon Atoms

Mildiomycin (**1-61**) is a peptidyl nucleoside antibiotic produced by *Streptoverticillium rimofaciens* that inhibits protein synthesis by blocking the peptidyltransferase site on the ribosomal subunit.[85,86] It is structurally similar to the antibiotics blasticidin S **1-62**, both sharing a common core structure.[87] Mildiomycin, however, contains a unique decose carbon skeleton. The identification and subsequent sequencing of the blasticidin S biosynthetic gene cluster in *Streptomyces griseochromogenes* (*bls*)[88] laid the foundation for the discovery of the mildiomycin biosynthetic gene cluster (*mil*).[89] Several initial steps regarding the biosynthesis of mildiomycin have been demonstrated. As shown in Figure 1.18, MilA, MilB and MilC have been demonstrated to convert cytidine monophosphate (CMP) to an early intermediate 5-hydroxymethyl-cytosylglucuronic acid (5-hydroxymethyl-

CGA, **1-63**).[88,89] MilG, a member of the radical SAM enzyme superfamily, is thought to be involved in the formation of the 2,3-deoxy-4-keto-2,3-ene-sugar moiety (**1-64**) of mildiomycin, since the *milG* disruption mutant accumulated 5-hydroxymethyl-CGA (**1-63**) in the fermentation broth. MilM and MilN are homologous to aspartate aminotransferase and dihydropicolinate synthetase, respectively. MilM is proposed to catalyze the deamination of arginine (**1-65**) to give **1-66**, which is subsequently coupled to the nucleoside by MilN (**1-64** + **1-66** $\rightarrow$ **1-67**, Figure 1.18).

The *S*-adenosyl-L-methionine analogues, sinefungin (**1-68**) and dehydrosinefungin (**1-69**), were first isolated from the culture media of *Streptomyces griseus* (Figure 1.19).[90] Both radiotracer feeding experiments and cell-free assays have provided preliminary clues regarding sinefungin biosynthesis.[91] It was proposed that adenosine and L-arginine are the biosynthetic precursors. However, the genes or enzymes involved in the process remain unidentified.

Tunicamycin (**1-70**) is a nucleoside antibiotic that was first isolated from culture media of *Streptomyces lysosuperificus* and subsequently *Streptomyces chartreusis* NRRL 3882.[92,93] Tunicamycin is extracted as a complex composed of closely related components having a general structure that includes an unusual 11-carbon aminodialdolose core, known as tunicamine (**1-71**), joined to *N*-acetylglucosamine (GlcNAc), uracil, and an amide-linked fatty acid (Figure 1.20).[94] Similar as uridyl lipopeptides (Section 1.2.1), tunicamycin disrupt the assembly of the bacterial cell wall via inhibition of MraY.[59]

The biosynthesis of the tunicamine core is only partially understood. Labeling studies with [2-^{14}C]-uridine, [6,6-$^{2}H_2$]-D-glucose, and D-[1-^{13}C]-glucose showed that the tunicamine core is derived from uridine (**1-72**) and UDP-GlcNAc (**1-73**).[95] The *S. chartreusis* genome was recently sequenced and a gene cluster for tunicamycin biosynthesis was identified (*tun*).[96,97] Within the gene cluster, TunB is identified as a putative member of the radical SAM enzyme superfamily. It was

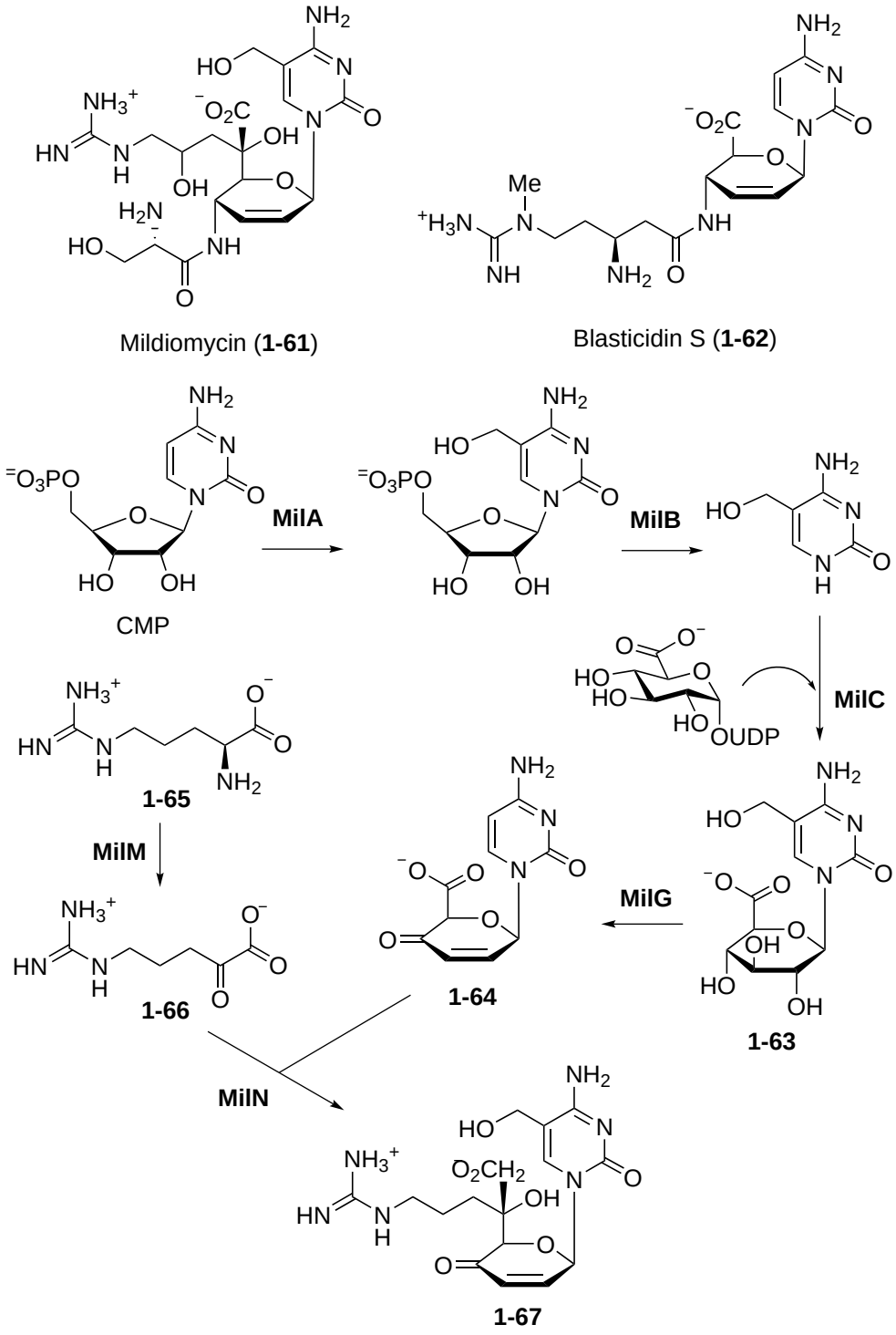
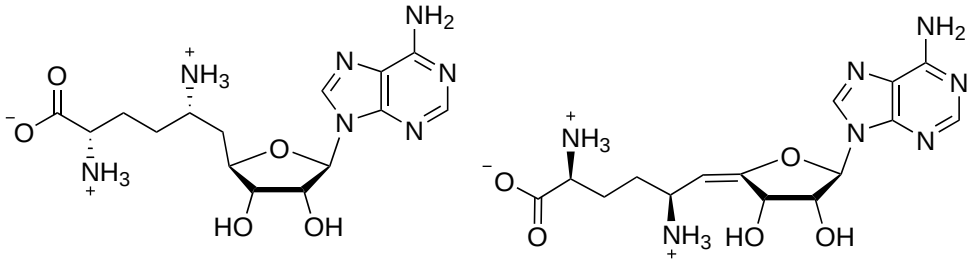


Figure 1.18: Proposed pathway for the biosynthesis of mildiomycin.



Sinefungin (**1-68**) Dehydrosinefungin (**1-69**)

Figure 1.19: Sinefungin and dehydrosinefungin.

then proposed that TunB catalyzes the condensation of uridine (**1-72**) with UDP-6-deoxy-GalNAc-5,6-ene (**1-76**) to yield UDP-*N*-acetyltunicamineuracil (**1-77**, Figure 1.20). This hypothesis is supported by the finding that a TunB knockout strain accumulates UDP-6-deoxy-GlcNAc-5,6-ene (**1-73**), which is believed to be the precursor of **1-76**).

The in vitro activities of purified recombinant proteins, TunA and TunF, were demonstrated after the discovery of gene cluster.[98] TunA is homologous to TDP-glucose 4,6-dehydratase, which carries out the conversion of TDP-glucose to TDP-4-keto-6-deoxy-glucose during the biosynthesis of 6-deoxyhexoses.[10,11] However, unlike a typical 4,6-dehydratase-catalyzed reaction involving an internal hydrogen transfer from C-4 to C-6, the TunA catalytic cycle is concluded by a 1,2-hydride addition to the 4-keto group of the 4-keto-6-deoxy-GlcNAc-5,6-ene intermediate (**1-74**), rather than 1,4-hydride addition to the adjacent C5–C6 double bond, to give UDP-6-deoxy-GlcNAc-5,6-ene (**1-75**) as the product (Figure 1.20).

Herbicidins are antibiotics with herbicidal activity produced by *Streptomyces sagamonensis*.[99,100] The herbicidins are adenosine nucleoside analogues that contain a tricyclic undecose scaffold. The array of herbicidins produced by a series of *S. sagamonensis* mutants led to a proposal that herbicidin A is derived from herbicidin F, of which the precursor is herbicidin G (**1-78**→**1-80**, Figure 1.21).[101] Hikizimycin (**1-81**), also named as anthelmycin, was isolated from the culture media

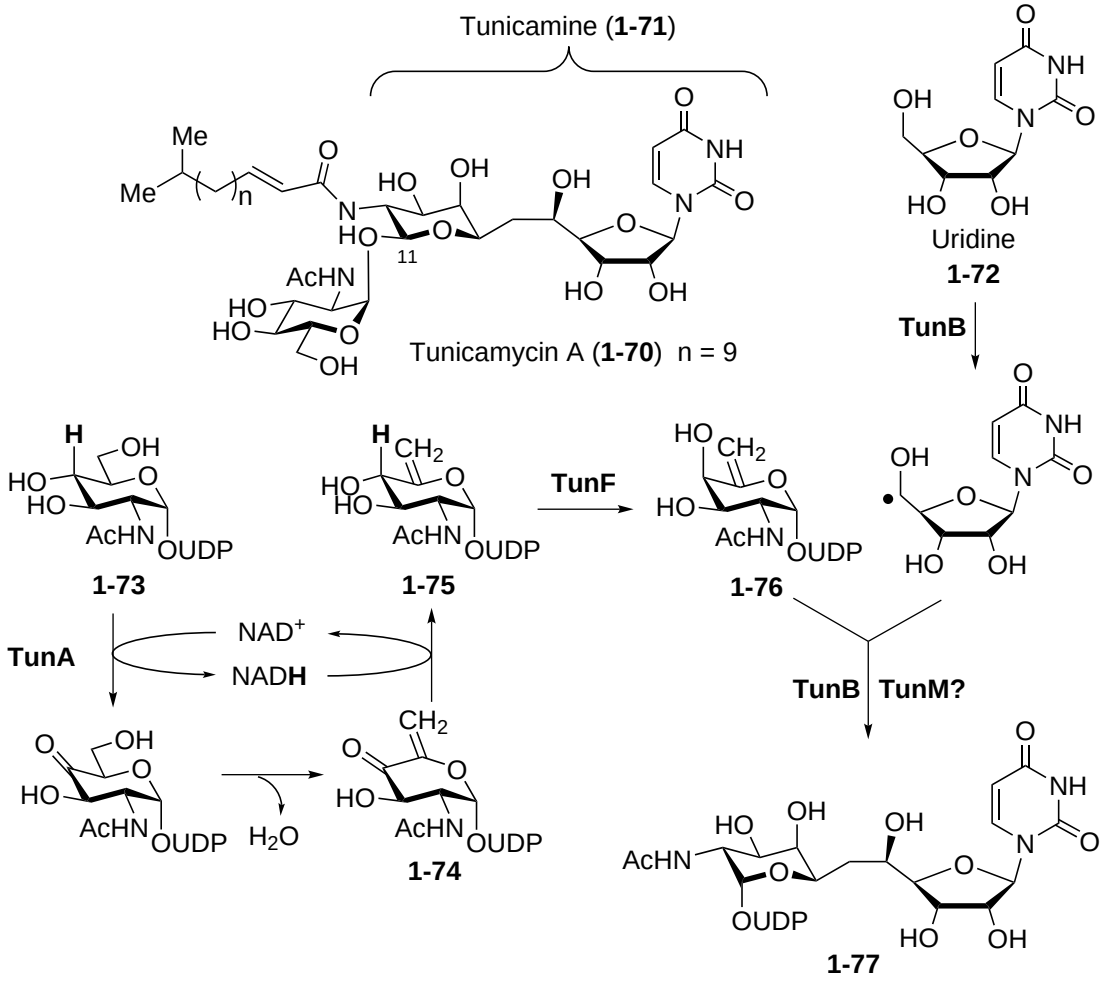


Figure 1.20: The proposed biosynthetic pathway of tunicamine core in tunicamycin A.

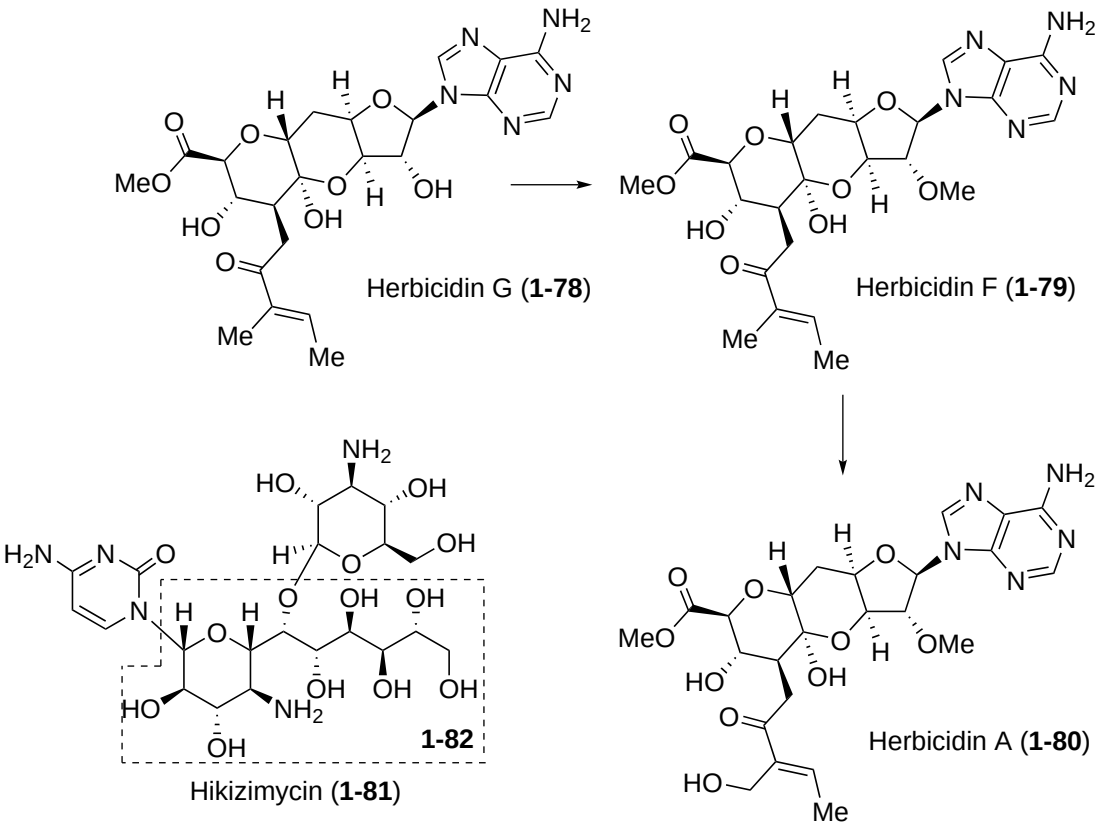


Figure 1.21: Hikizimycin and the proposed biosynthetic pathway of herbicidins.

of *Streptomyces* sp. strain A-5 and exhibits antifungal activity.[102] Hikizimycin is a cytosine nucleoside analogue containing hikosamine (**1-82**), a 4-amino-4-deoxy-undecose sugar (Figure 1.21).[103,104]

1.3 Summary and Thesis Statement

The intricate chemical structures of natural products has inspired research to study their mode of actions, structure-activity relationships and biosynthesis. Development of methods and strategies for the total synthesis of natural products remains one of the most important fields in organic chemistry.[105] Investigating the construction of the elaborately structured natural products by enzymatic catalysis is also an important topic in enzymology. The accumulated knowledge has allowed pathway engineering and synthetic biology to become useful tools for drug development[106] and biofuel research.[107] Moreover, the strategy combining the information of natural product biosynthesis and the capability of organic synthesis serves as a powerful approach used in modern pharmaceutical research. A recent success in the production of antimalarial artemisinin fully demonstrated the potential of cross-field integration.[108] The rarity of thio- and high-carbon sugars does not prevent them being used as important therapeutic agents. However, knowledge about how thiosugars and high-carbon sugars are biosynthesized sugars mostly remains elusive due to the scarcity of their natural occurrence. Thus, the focus of this study is to investigate the biosynthesis of the natural products possessing these two structural features.

Lincomycin A (**1-17**) is a clinically important antibiotic noted for its unusual C-1 methylthio-substituted octose moiety. Based on previous feeding experiments and sequence analysis, a proposed biosynthetic pathway of lincomycin has been proposed and is described in Chapter 2. LmbR (transaldolase) was shown to catalyze the condensation of D-ribose 5-phosphate and D-fructose 6-phosphate or D-

sedoheptulose 7-phosphate. The resulting octulose 8-phosphate was subsequently was isomerized to octose 8-phosphate through the catalysis of LmbN (isomerase). Chemical synthesis of the expected product was carried out and the products were used for the structural confirmation of LmbN-product. In Chapter 3, the identification of GDP-octose as a key intermediate in the biosynthesis of lincomycin A is described. The conversion of octose 1,8-bisphosphate intermediate to octose 1-phosphate by LmbK (phosphatase) was demonstrated. The following C-1 activation of octose 1-phosphate catalyzed by LmbO (guanylyltransferase) resulted in the formation of GDP-octose.

Chapter 4 focuses on the characterization of enzymes involved in GDP-octose modifications, including C-4 epimerization, C-8 dehydration and C-6 amination. Overexpression and purification of the proteins LmbM (epimerase), LmbL (dehydrogenase), LmbZ (oxidoreductase) and LmbS (aminotrasnferase) allowed further investigation of their proposed functions. Finally, in Chapter 5, sulfur incorporation was pursued using the GDP-aminosugar with a series of sulfur donors and the putative *S*-glycosyltransferase, LmbT. Moreover, a unique condensation system resembling nonribosomal peptide synthetase was shown to catalyze the amide bond formation between the proline derivative and methylthiolincosamide. The final key step of lincomcyin A is the C–S bond cleavage to yield C-1 free thiol group. The pyridoxal 5′-phosphate-dependent enzyme, LmbF, was purified. A mechanistic proposal regarding the bifurcation point of the lincomcyin A and celesticetin biosynthetic pathways is also described. These results represent the first characterization of the biosynthesis of a thiooctose-containing natural product.

Chapter 2

Biosynthetic Studies of Lincomycin A (I): Identification of the Octose 8-Phosphate as the Biosynthetic Intermediate*

2.1 Introduction

Lincomycin A (**2-1**) is a clinically important antibiotic against Gram-positive bacteria.[109] It was isolated from a soil actinomycete, *Streptomyces lincolnensis*, which was named after the city of Lincoln in Nebraska where the strain was discovered.[28] Similar to macrolide antibiotics, lincomycin interacts with peptidyltransferase domain of the 50S ribosomal subunit due to its structural resemblance to the 3′-end of L-Pro-Met-tRNA and deacetylated-tRNA. This, in turn, inhibits the bacterial protein synthesis.[30,110] Nowadays, lincomycin A, clindamycin (**2-8**) and pirlimycin (**2-9**) are three lincosamide compounds widely used to treat *Staphylococcal* and *Streptococcal* infections as both human and veterinary medicine (Figure 2.1B). Both clindamycin (**2-8**) and pirlimycin (**2-9**) are semisynthetic derivatives of lincomycin A (**2-1**).[111,112]

Resistance to lincosamides has been observed in several strains of *Staphylococcus* due to the SAM-dependent methylation of 23S rRNA. This modification greatly decreases the affinity between the ribosome and the antibiotics, which leads to coresistance to macrolide and streptogramin B-type compounds.[113] The

*Partial content of this chapter was published as "Construction of the Octose 8-Phosphate Intermediate in Lincomycin A Biosynthesis: Characterization of the Reactions Catalyzed by LmbR and LmbN." Sasaki, E.; Lin, C.-I; Lin, K.-Y.; Liu, H.-w., *J. Am. Chem. Soc.*, **2012**, *134*, 17432–17435. The experiments were designed by E. S. and performed by E. S. and C.-I L. One of the enzymes was prepared by E. S. with the help from K.-Y. Lin. This research project was supervised by H.-w. Liu.

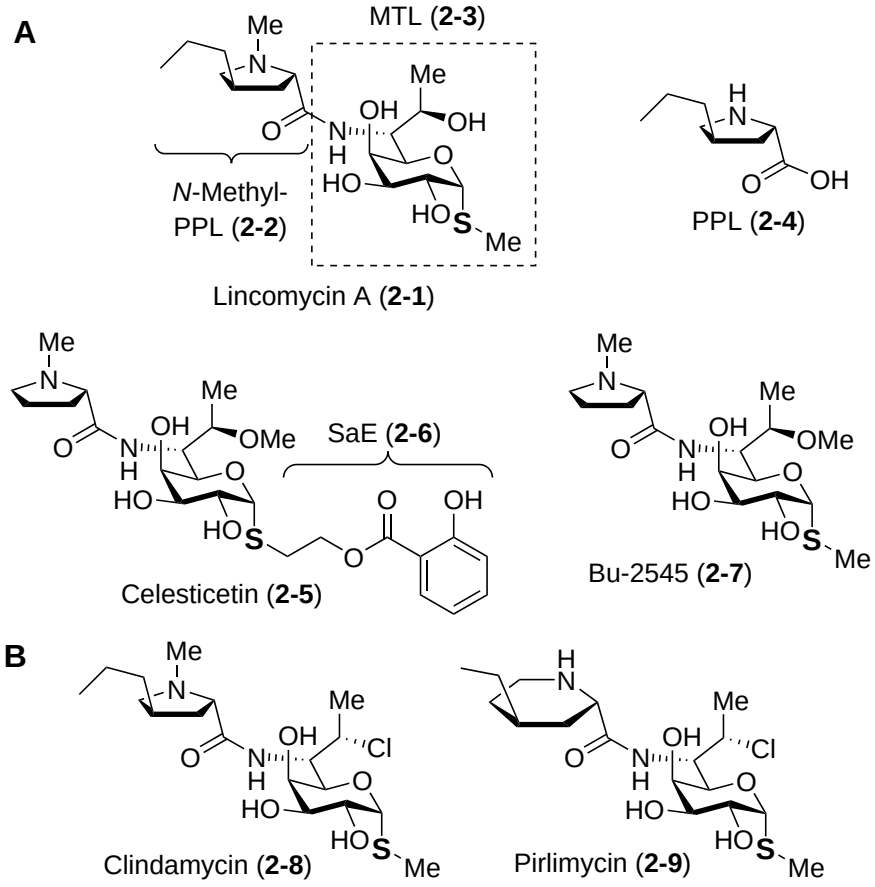


Figure 2.1: (A) Naturally occurring lincosamides. (B) Semisynthetic lincosamides from lincomycin A.

other type of resistance to lincomycin is more specific and is due to adenylation of the C3′-hydroxyl group of the sugar moiety of lincomycin A. This nucleotidyltransfer reaction is catalyzed by an *O*-nucleotidyltransferase, encoded by the gene *linA*.[114]

The structure of lincomycin A consists of an amino acid derivative, *N*-methyl-4-propyl-L-proline (*N*-methyl-PPL, **2-2**), and an unusual methylthiolincosamide (MTL, **2-3**) moiety.[31] Similar thiooctoses are also found in several antimicrobial agents, such as celesticetin (**2-5**) from *Streptomyces caelestis*[33] and Bu-2545 (**2-7**) from *Streptomyces* strain H230-5[32] (Figure 2.1A). The biosynthetic pathway of lincomycin A is believed to proceed through a heterogeneously rooting biphasic pathway producing PPL (**2-4**) and MTL (**2-3**) separately.[115] The two units are later condensed and methylated to yield lincomycin A.

About thirty years ago, feeding experiments were conducted in order to identify the biosynthetic origin of the octose sugar in lincomycin A (**2-1**). When *S. lincolnensis* was grown with a blend of unlabeled and [$^{13}C_6$]-glucose as the sole carbon source, the ^{13}C labeling pattern of the resulting MTL (**2-3**) suggested that MTL is formed via the condensation of a three-carbon (C_3) and a five-carbon (C_5) unit derived from glucose.[116] This led to a proposal that the C_5 unit might be an intermediate from the pentose phosphate pathway and that the C_3 unit might be derived from D-sedoheptulose 7-phosphate (S7P, **2-11**, Figure 2.2B). These two units may be joined together by a transaldolase-catalyzed reaction analogous to that involved in the pentose phosphate pathway (Figure 2.2A).

The genes required for lincomycin A biosynthesis were later identified and sequenced from *S. lincolnensis* strains 78-11[115] and ATCC 25466 (designated as *lmb*).[117] When *Streptomyces coelicolar* was transformed with a cosmid containing the *lmb* cluster from *S. lincolnensis* ATCC 25466, heterologous lincomycin A production was observed.[117] Hybridization analysis using probes designed according to the

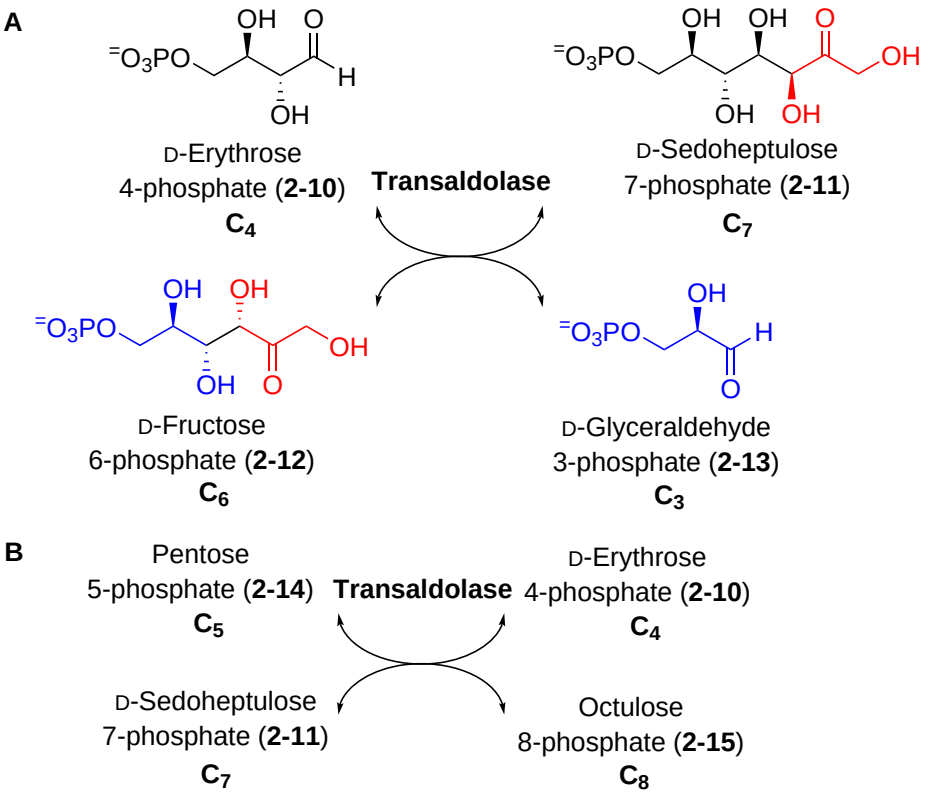


Figure 2.2: (A) Transaldolase reaction in pentose phosphate pathway. (B) Proposed transaldolase reaction MTL biosynthesis.

lincomcyin A biosynthetic gene cluster allowed the identification of the celesticetin biosynthetic gene cluster in *S. caelestis* ATCC 1584 (designated as *ccb*).[118] Interestingly, both gene clusters encode a putative transaldolase (LmbR and CcbR) that may catalyze an aldol condensation to produce the octose sugar precursor of MTL (Figure 2.2B). Moreover, several genes, *lmbO*, *lmbM* and *lmbS*, showing sequence homology to those found in various NDP-deoxyhexose pathways also exist in both *lmb* and *ccb* clusters. This finding led to the hypothesis that an NDP-activated octopyranose is a possible intermediate in the MTL biosynthetic pathway.[115]

In this chapter, the genes in both lincomycin A and celesticetin gene clusters are annotated using the updated database. The proposed biosynthetic pathway of the thiooctose moiety is build based on the sequence similarity of each gene product to known enzymes reported in literature. The enzymatic function of the putative transaldolase, LmbR, is characterized and the biosynthetic origin of octose is elucidated. This validation is crucial for establishing the entire pathway. Another enzyme predicted to involve in the early state of lincosamide biosynthesis is an isomerase, LmbN. Investigation and stereochemical analysis of LmbN-catalyzed reaction product is also presented in this chapter.

2.2 Experimental Procedures

2.2.1 General

Materials

All chemicals and reagents were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA) except for D-sedoheptulose 7-phosphate from Carbosynth (San Diego, CA), and were used without further purification unless otherwise specified. Enzymes and molecular weight standards used for the cloning experiments were obtained from Invitrogen (Carlsbad, CA) or New England Biolabs (Ipswich, MA). Reagents for sodium dodecyl sulfate poly-

acrylamide gel electrophoresis (SDS-PAGE) were purchased from Bio-Rad (Hercules, CA), with the exception of the protein molecular weight markers, which were obtained from Invitrogen. Analytical C_{18} HPLC columns were products of Varian (Palo Alto, CA). Amicon YM-10 ultrafiltration membranes were bought from Millipore (Billerica, MA).

Instrumentation

HPLC was performed on a Beckman Coulter System Gold HPLC equipped with a UV detector or a Corona CAD (charged aerosol detector, ESA Biosciences, Chelmsford, MA). Analytical thin layer chromatography (TLC) was carried out on precoated TLC glass plates (silica gel, grade 60, F254, 0.25 mm layer thickness) purchased from EMD chemicals (Madison, WI). Flash column chromatography was performed on silica gel (40–63 μm, 60 Å) obtained from SiliCycle Inc (Quebec City, Canada). NMR spectra were acquired using a Varian Unity 400 MHz, 500 MHz or 600 MHz spectrometer, and chemical shifts (in ppm) are reported relative to that of the solvent peak (δ_H = 7.24 and δ_C = 77.0 for deuterated chloroform, δ_H = 4.67 for deuterium oxide). Mass spectra were recorded at the Mass Spectrometry core facility in the Department of Chemistry and the Proteomics Facility in College of Pharmacy at the University of Texas, Austin. Optical rotations were measured using an ATAGO AP-300 automatic polarimeter at a path length of 1 dm.

2.2.2 Sequence Analysis and Function Prediction of Proteins

The primary amino acid sequence of each protein encoded in the lincomcyin A biosynthetic gene cluster from *S. lincolnensis* ATCC 25466 (accession number EU124663) and the celesticetin biosynthetic gene cluster from *S. caelestic* ATCC 1584 were analyzed using The Universal Protein Resource (UniProt) database by the basic local alignment search (BLAST).

2.2.3 ESI-MS Analysis of LmbR

D-Sedoheptulose 7-phosphate was used as substrate in the LmbR reaction. The purified *C*-His_6-LmbR protein stock solution was added to 50 mM Tris · HCl buffer (pH 8.0) (final concentration of protein was 40 μM) and incubated with D-sedoheptulose 7-phosphate (S7P, **2-11**, 10 mM) at 30 °C for 15 min. The resulting solution (200 μL) was cooled on ice. To this solution was added 20 μL of cold $NaBH_4$ solution in methanol (500 mM; final concentration of $NaBH_4$ was 45 mM). The mixture was incubated on ice for 1 h. The excess salts and glycerol were removed at 4 °C using an Amicon ultrafiltration unit through a YM-10 membrane. The resulting protein solution was subjected to ESI-MS analysis. As a control, the purified *C*-His_6-LmbR protein (40 μM) in 50 mM Tris · HCl buffer (pH 8.0) in the absence of S7P was also treated with $NaBH_4$ and subjected to ESI-MS analysis. The recombinant LmbR was prepared by another graduate student, Eita Sasaki.

2.2.4 Synthesis of 1,2,3,4,6,7,8-Hepta-*O*-acetyl-α,β-D-*erythro*-D-*gluco*-octopyra-nose (2-27)

As shown in Figure 2.3, compound **2-27** was synthesized from methyl α-D-glucopyranose (**2-16**) following a previously reported procedure with minor modifications.[119]

Methyl 2,3,4-tri-*O*-benzyl-α-D-*gluco*-pyranoside (2-19).
Compound **2-16** (18.05 g, 93.0 mmol) and 4-dimethylaminopyridine (DMAP, 0.56 g, 4.58 mmol) were dissolved in anhydrous DMF (150 mL) and the reaction mixture was cooled to 0 °C. The reaction mixture was then mixed with triethylamine (45 mL, 320 mmol) and stirred at 0 °C for 10 min. To this solution was added TBSCl (15.73 g, 104.4 mmol) portion-wise and the reaction mixture was brought to room temperature slowly with stirring. After 16 h, the reaction mixture was quenched by adding water and concentrated *in vacuo*. The resulting residue was dissolved

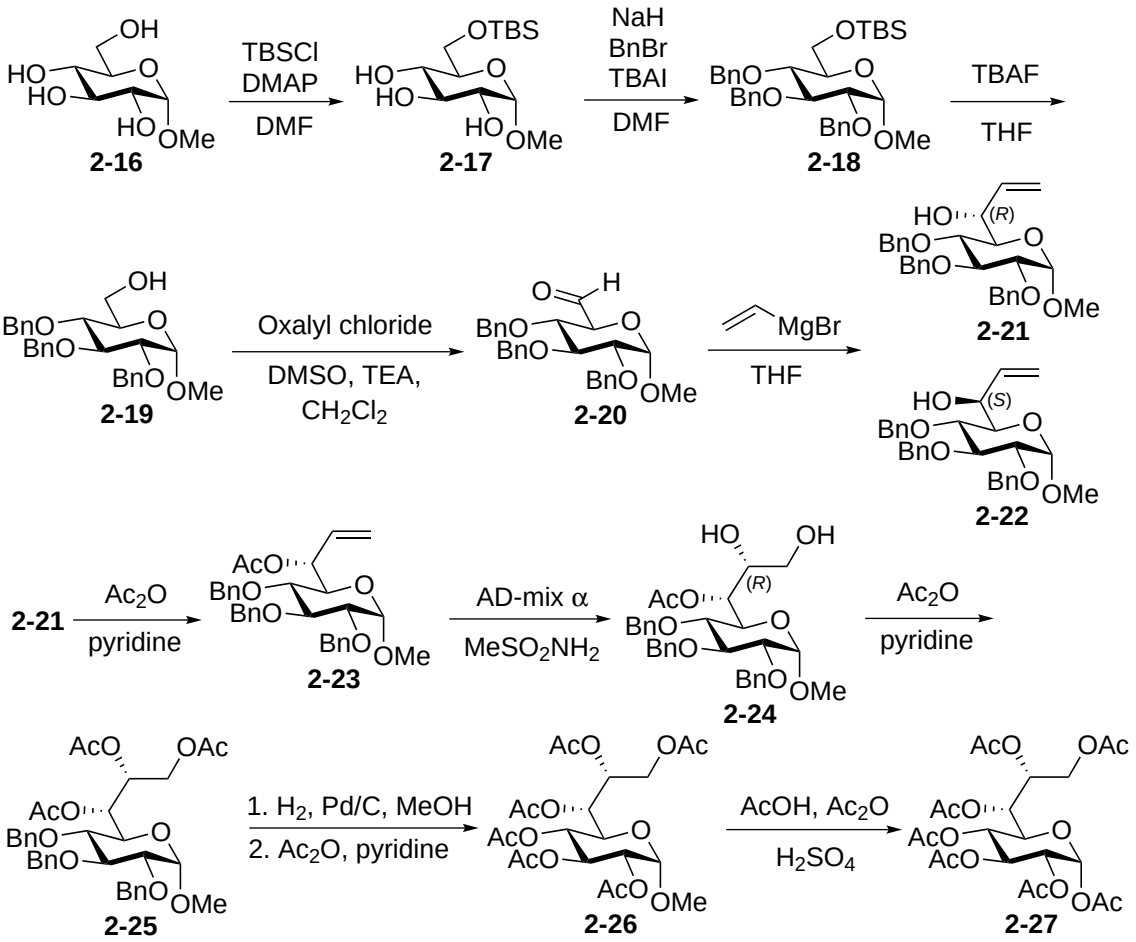


Figure 2.3: Synthetic scheme of 1,2,3,4,6,7,8-hepta-*O*-acetyl-α,β-D-*erythro*-D-*gluco*-octopyranose (**2-27**).

in ethyl acetate. The organic solution was washed with water (1×) and brine (1×), and then dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation under reduced pressure, the crude product was purified by flash chromatography (silica gel; EtOAc only) to yield **2-17** ($C_{13}H_{28}O_6Si$, 22.64 g, 83%) as a white solid. TLC (EtOAc only) R_f = 0.18.

Compound **2-17** (13.32 g, 43.18 mmol) and tetrabutylammonium iodide (TBAI, 0.8 g, 4.6 mmol) were dissolved in anhydrous DMF (150 mL) and cooled to 0 °C. The solution was then treated with sodium hydride (60%, 9.20 g, 230.0 mmol) portionwise and stirred at 0 °C for 15 min. Benzyl bromide (21.0 mL, 176 mmol) was added slowly and the reaction mixture was brought to room temperature gradually. After 16 h, the reaction mixture was quenched by adding saturated ammonium chloride solution and concentrated *in vacuo*. The resulting residue was mixed with water and extracted with ethyl ether (3×). The combined organic layers were dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation under reduced pressure, the crude product was purified using flash chromatography (silica gel; EtOAc/hexanes (1:19→1:9)) to yield **2-18** ($C_{34}H_{46}O_6Si$, 7.55 g, 30%) as a white solid; TLC (EtOAc/hexanes (1:9)) R_f = 0.27.

Compound **2-18** (7.55 g, 13.0 mmol) was dissolved in anhydrous THF (200 mL) and added tetrabutylammonium fluoride solution (1.0 M in THF, 40 mmol). The resulting mixture was stirred at room temperature for 16 h. The reaction was quenched by adding saturated ammonium chloride solution and extracted with ethyl acetate (3×). The combined organic layers were dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation under reduced pressure, the crude product was purified using flash chromatography (silica gel; EtOAc/hexanes (1:1)) to yield **2-19** ($C_{28}H_{32}O_6$, 4.57 g, 75%) as a colorless solid. TLC (EtOAc/hexanes (1:9)) R_f = 0.08. ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) 7.37-7.26 (15 H, m, Bn×3), 4.98 (1 H, d, J = 11.2 Hz, CH_2 of Bn), 4.87 (1 H, d, J = 11.2 Hz, CH_2 of Bn), 4.83 (1

H, d, J = 11.2 Hz, CH_2 of Bn), 4.79 (1 H, d, J = 12.4 Hz, CH_2 of Bn), 4.65 (1 H, d, J = 12.4 Hz, CH_2 of Bn), 4.63 (1 H, d, J = 11.2 Hz, CH_2 of Bn), 4.56 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.00 (1 H, dd, $J_{3,4}$ = 9.2 Hz, $J_{4,5}$ = 9.2 Hz, H4), 3.77–3.62 (3 H, m, H5, H6$_a$, H6$_b$), 3.51 (1 H, dd, $J_{2,3}$ = 9.6 Hz, $J_{3,4}$ = 9.2 Hz, H3), 3.49 (1 H, d, $J_{1,2}$ = 3.6 Hz, $J_{2,3}$ = 9.6 Hz, H2), 3.35 (3 H, s, OMe). The ^{1}H NMR spectral data of **2-19** are consistent with those previously reported.[119]

Methyl 2,3,4-tri-*O*-benzyl-α-D-*gluco*-hexodialdo-1,5-pyranoside (2-20).

Under argon, a mixture of anhydrous dichloromethane (15 mL) and DMSO (2.9 mL, 37.5 mmol) was cooled to −78 °C in dry ice-acetone bath and then added oxalyl chloride (1.6 mL, 18.3 mmol) dropwise. The reaction mixture was stirred at −78 °C for 10 min and then added anhydrous dichloromethane solution of alcohol **2-19** (4.28 g, 9.22 mmol) dropwise. The reaction mixture was stirred at −78 °C for 1 h. At −78 °C, the reaction mixture was added triethylamine (6.5 mL, 46.6 mmol) slowly. After addition, the reaction was brought to 0 °C in an ice-water bath and stirred for an additional 30 min. The reaction was then diluted with dichloromethane and quenched by adding water. The resulting mixture was extracted with dichloromethane (3×) and the combined organic layers were dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation, the crude aldehyde **2-20** was used for the next step without further purification.

Methyl 2,3,4-tri-*O*-benzyl-7,8-dideoxy-D-*glycero*-α-D-*gluco*-oct-7-enopyranoside (2-21).[120]

Aldehyde **2-20** (∼9.22 mmol) was generated *in situ* as described above and dissolved in anhydrous THF (90 mL). The aldehyde solution was cooled to 0 °C and then added vinylmagnesium bromide solution (0.7 M in THF, 37 mL) slowly. The reaction mixture was stirred at room temperature for 16 h. It was then diluted with ethyl acetate and quenched by adding water. The reaction mixture was extracted with ethyl acetate (3×) and the combined organic layers were dried over anhy-

drous Na_2SO_4. After the volatiles were removed by rotary evaporation, the crude product was purified using flash chromatography (silica gel; EtOAc/hexanes (1:9 $\rightarrow$ 1:4)) to elute the undesired L-*glycero*-α-D-*gluco*-isomer **2-22** (1.56 g, 34%). The desired product **2-21** ($C_{30}H_{34}O_6$, 0.92 g) was then eluted by EtOAc/hexanes (1:2) to give a white solid in 20% yield; TLC (EtOAc/hexanes (1:1)) R_f = 0.48. ^{1}H NMR ($CDCl_3$, 400 MHz) of **2-21** δ (ppm) 7.36-7.25 (15 H, m, Bn×3), 5.90 (1 H, ddd, $J_{6,7}$ = 6.8 Hz, $J_{7,8cis}$ =10.8 Hz, $J_{7,8trans}$ = 17.2 Hz, H7), 5.25 (1 H, ddd, $J_{6,8trans}$ = $J_{8cis,8trans}$ = 1.6 Hz, $J_{7,8trans}$ = 17.2 Hz, $H8_{trans}$), 5.19 (1 H, ddd, $J_{6,8cis}$ = $J_{8cis,8trans}$ = 1.6 Hz, $J_{7,8cis}$ = 10.8 Hz, $H8_{cis}$), 5.00 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.92 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.78 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.78 (1 H, J = 12 Hz, CH_2 of Bn), 4.65 (1 H, J = 12 Hz, CH_2 of Bn), 4.60 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.59 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.34 (1 H, m, H6), 4.02 (1 H, dd, $J_{2,3}$ = $J_{3,4}$ = 9.2 Hz, H3), 3.79 (1 H, dd, $J_{5,6}$ = 3.6 Hz, $J_{4,5}$ = 10.4 Hz, H5), 3.48 (1 H, dd, $J_{2,3}$ = 9.2 Hz, $J_{1,2}$ = 3.6 Hz, H2), 3.43 (1 H, dd, $J_{3,4}$ = 9.2 Hz, $J_{4,5}$ = 10.4 Hz, H4), 3.38 (3 H, s, OMe). The ^{1}H NMR spectral data of **2-21** are consistent with those previously reported.[120] The configuration at the C-6 position of **2-21** was determined by Moshers ester analysis as described in Section 2.2.6.

The undesired L-*glycero*-α-D-*gluco*-isomer **2-22**: 7.38-7.26 (15 H, m, Bn×3), 5.95 (1 H, ddd, $J_{6,7}$ = 4.8 Hz, $J_{7,8cis}$ =10.4 Hz, $J_{7,8trans}$ = 17.2 Hz, H7), 5.32 (1 H, ddd, $J_{6,8trans}$ = $J_{8cis,8trans}$ = 1.6 Hz, $J_{7,8trans}$ = 17.2 Hz, $H8_{trans}$), 5.19 (1 H, ddd, $J_{6,8cis}$ = $J_{8cis,8trans}$ = 1.6 Hz, $J_{7,8cis}$ = 10.4 Hz, $H8_{cis}$), 5.00 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.94 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.85 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.80 (1 H, J = 12.4 Hz, CH_2 of Bn), 4.70 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.65 (1 H, J = 12.4 Hz, CH_2 of Bn), 4.57 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.38 (1 H, m, H6), 4.02 (1 H, dd, $J_{2,3}$ = $J_{3,4}$ = 9.2 Hz, H3), 3.70 (1 H, dd, $J_{3,4}$ = 9.2 Hz, $J_{4,5}$ = 10.0 Hz, H4), 3.61 (1 H, dd, $J_{5,6}$ = 1.6 Hz, $J_{4,5}$ = 10.0 Hz, H5), 3.51 (1 H, dd, $J_{2,3}$ = 9.2 Hz, $J_{1,2}$ = 3.6 Hz, H2), 3.19 (3 H, s, OMe).

Methyl 6-*O*-acetyl-2,3,4-tri-*O*-benzyl-7,8-dideoxy-D-*glycero*-α-D-*gluco*-oct-7-eno

-pyranoside (2-23).

Alcohol **2-21** (1.75 g, 3.57 mmol) was dissolved in anhydrous pyridine (9 mL) and added acetic acid (9 mL). The reaction mixture was stirred at room temperature for 16 h. The reaction was quenched by adding methanol slowly and then concentrated *in vacuo*. The resulting residue was dissolved in ethyl acetate. The organic solution was washed with brine (1×) and then dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation under reduced pressure, the crude product was purified using flash chromatography (silica gel; EtOAc/hexanes (1:4)) to give **2-23** ($C_{32}H_{36}O_7$, 1.19 g, 63%) as a light yellow syrup. TLC (EtOAc / hexanes (1:1)) R_f = 0.32. ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) 7.39-7.28 (15 H, m, Bn×3), 5.86 (1 H, ddd, $J_{6,7}$ = 7.6 Hz, $J_{7,8cis}$ =10.8 Hz, $J_{7,8trans}$ = 17.2 Hz, H7), 5.66-5.64 (1 H, m, H6), 5.29 (1 H, ddd, $J_{6,8trans}$ = $J_{8cis,8trans}$ = 1.2 Hz, $J_{7,8cis}$ = 10.8 Hz, $H8_{cis}$), 5.24 (1 H, ddd, $J_{6,8cis}$ = $J_{8cis,8trans}$ = 1.2 Hz, $J_{7,8trans}$ = 17.2 Hz, $H8_{trans}$), 5.01 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.93 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.81 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.80 (1 H, J = 12 Hz, CH_2 of Bn), 4.68 (1 H, J = 12 Hz, CH_2 of Bn), 4.65 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.65 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.05 (1 H, dd, $J_{2,3}$ = 9.6 Hz, $J_{3,4}$ = 8.8 Hz, H3), 3.91 (1 H, dd, $J_{5,6}$ = 1.6 Hz, $J_{4,5}$ = 10 Hz, H5), 3.50 (1 H, dd, $J_{2,3}$ = 9.6 Hz, $J_{1,2}$ = 3.6 Hz, H2), 3.41 (1 H, dd, $J_{3,4}$ = 8.8 Hz, $J_{4,5}$ = 10.4 Hz, H4), 3.41 (3 H, s, OMe), 2.07 (3 H, s, OAc). ^{13}C NMR ($CDCl_3$, 100 MHz) δ (ppm) 169.6, 138.4, 137.89, 137.86, 131.3, 128.34, 128.29, 127.94, 127.85, 127.63, 127.55, 120.47, 97.6, 82.0, 79.8, 77.7, 75.7, 74.4, 73.7, 73.2, 71.5, 54.9, 21.0. ESI-HRMS Calcd. for $C_{32}H_{36}O_7[M + Na]^+$ 555.2359, Found: 555.2354.

Methyl 6-*O*-acetyl-2,3,4-tri-*O*-benzyl-α-D-*erythro*-D-*gluco*-octopyranoside (2-24). Compound **2-23** (1.19 g, 2.23 mmol) was suspended in 20 mL of co-solvent ^{t}BuOH : H_2O (1:1). The mixture was cooled to 4 °C in cold room and added AD-mix α (3.31 g) and methanesulfonamide (0.22 g, 2.31 mmol). The reaction mixture was stirred at 4 °C until the starting material was all consumed as judged by TLC analysis. The

reaction was quenched by adding sodium metabissulfite and then concentrated *in vacuo*. The resulting residue was dissolved in ethyl acetate. The organic solution was washed with water (1×) and brine (1×), and then dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation, the crude diol **2-24** was collected and used for the next step without further purification.

Methyl 6,7,8-tri-*O*-acetyl-2,3,4-tri-*O*-benzyl-α-D-*erythro*-D-*gluco*-octopyranoside (2-25).

The crude diol **2-24** (2.23 mmol) was dissolved in anhydrous pyridine (9 mL) and added acetic anhydride (9 mL). The reaction mixture was stirred at room temperature for 16 h. The reaction was quenched by adding ethanol slowly and then concentrated in vacuo. The resulting residue was dissolved in ethyl acetate. The organic solution was washed with brine (1×) and then dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation under reduced pressure, the crude **2-25** was used for the later step without further purification.

1,2,3,4,6,7,8-Hepta-*O*-acetyl-α,β-D-*erythro*-D-*gluco*-octopyranose (2-27).

A methanol solution (5 mL) of compound **2-25** (0.30 g, 0.46 mmol) and 5% Pd/C (50 mg) was stirred under hydrogen atmosphere (1 atm) for 16 h. The reaction mixture was filtered through Celite and washed with methanol. The filtrate was concentrated and re-dissolved in anhydrous pyridine (5 mL). The solution was then added acetic anhydride (5 mL). The reaction mixture was stirred at room temperature for 16 h and concentrated to yield **2-26**.

Compound **2-26** (0.10 g, 0.2 mmol) was dissolved in acetic acid (3 mL) and added acetic anhydride (3 mL). The reaction mixture was cooled to 0 °C in an ice-water bath and then added concentrated sulfuric acid (0.3 mL) dropwise. After stirring at 4 °C in the cold room for 16 h, the reaction mixture was diluted with dichloromethane and then quenched by adding saturated sodium bicarbonate solution slowly. The resulting solution was extracted with dichloromethane (3). The combined organic

layers were washed with water (1×) and brine (1×), and then dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation, the crude product was purified using flash chromatography (silica gel; EtOAc/hexanes (1:1)) to yield **2-27** ($C_{22}H_{30}O_{15}$, 75 mg, 71%, $>$ 99.5% α anomer based on the ^{1}H NMR) as a white solid; TLC (EtOAc/hexane (1:1)) R_f = 0.26. ^{1}H NMR ($CDCl_3$, 500 MHz) δ (ppm) 6.28 (1 H, d, $J_{1,2}$ = 3.5 Hz, H1), 5.37 (1 H, dd, $J_{2,3}$ = 10.5 Hz, $J_{3,4}$ = 9.0 Hz, H3), 5.27 (1 H, dd, $J_{5,6}$ = 3.5 Hz, $J_{6,7}$ = 6.0 Hz, H6), 5.16 (1 H, dd, $J_{3,4}$ = 9.0 Hz, $J_{4,5}$ = 10.5 Hz, H4), 5.17–5.14 (1 H, m, H7), 4.99 (1 H, dd, $J_{1,2}$ = 3.5 Hz, $J_{2,3}$ = 10.5 Hz, H2), 4.31 (1 H, dd, $J_{7,8a}$ = 2.5 Hz, $J_{8a,8b}$ = 12.5 Hz, H8a), 4.15 (1 H, dd, $J_{4,5}$ = 10.5 Hz, $J_{5,6}$ = 3.5 Hz, H5), 4.10 (1 H, dd, $J_{7,8b}$ = 6.0 Hz, $J_{8a,8b}$ = 12.5 Hz, H8b), 2.14 (3 H, s, OAc), 2.06 (3 H, s, OAc), 2.05 (3 H, s, OAc), 2.03 (3 H, s, OAc), 2.01 (3 H, s, OAc), 1.98 (3 H, s, OAc), 1.97 (3 H, s, OAc). ^{13}C NMR ($CDCl_3$, 125 MHz) δ (ppm) 170.6, 170.2, 169.7, 169.6, 169.4, 169.3, 168.5, 88.7, 70.6, 70.1, 69.6, 69.5, 68.9, 68.8, 61.8, 20.8, 20.7, 20.64 (×2), 20.59 (×2), 20.4. ESI-HRMS Calcd. for $C_{22}H_{30}O_{15}[M + Na]^+$ 557.1482, Found: 557.1474.

2.2.5 Synthesis of 1,2,3,4,6,7,8-Hepta-*O*-acetyl-α,β-D-*erythro*-D-*galacto*-octopyranose (2-39)

As shown in Figure 2.4, compound **2-39** was synthesized from methyl α-D-galactopyranose (**2-28**) following a previously reported procedure with minor modifications.[121]

Methyl 2,3,4-tri-*O*-benzyl-α-D-*galacto*-pyranoside (2-31).
Compound **2-28** (9.02 g, 42.5 mmol) was dissolved in anhydrous DMF (100 mL) and the solution was cooled to 0 °C in an ice-water bath. To this solution was added imidazole (7.27 g, 107 mmol) and the resulting mixture was stirred at 0 °C for 5 min. TBSCl (7.84 g, 52.0 mmol) was added portion-wise and the reaction mixture was brought to room temperature gradually. After 16 h, the reaction

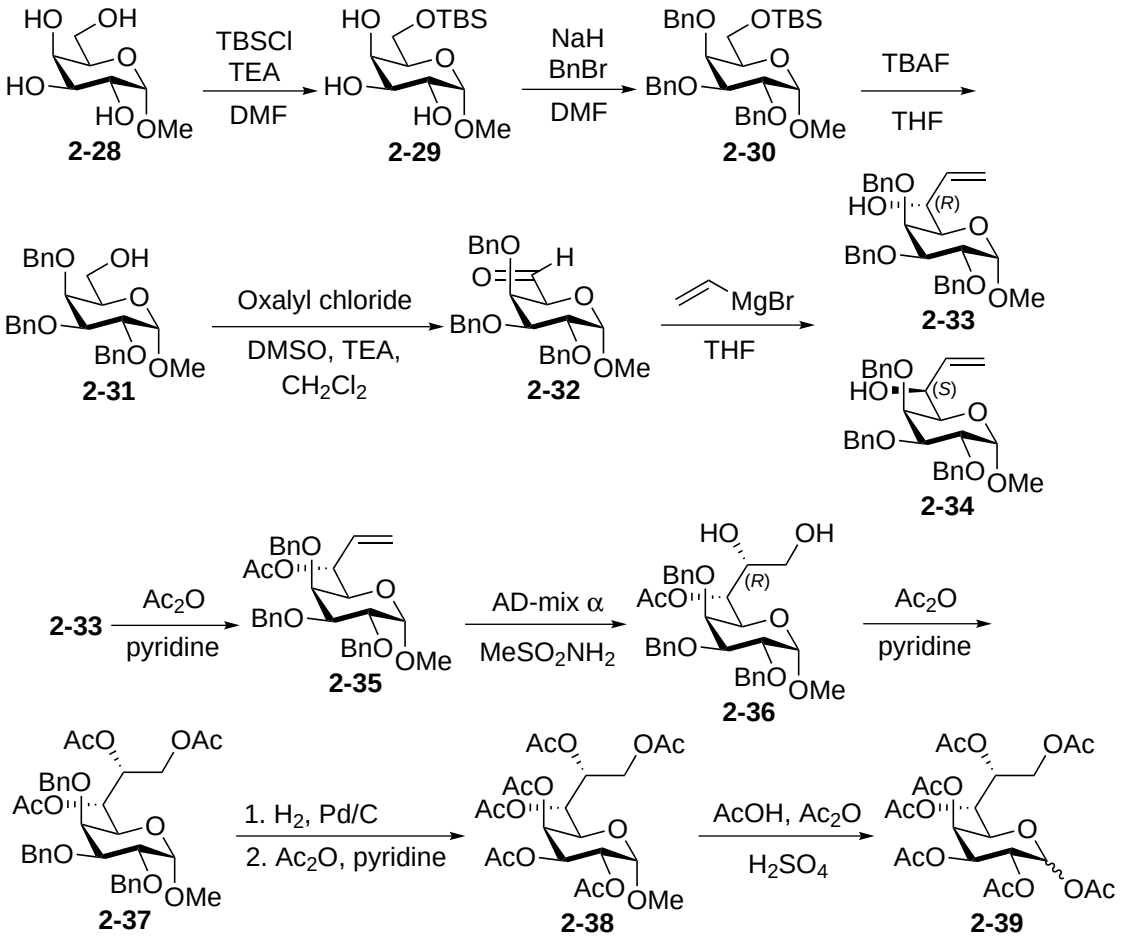


Figure 2.4: Synthetic scheme of 1,2,3,4,6,7,8-hepta-*O*-acetyl-α,β-D-*erythro*-D-*galacto*-octopyranose (**2-39**).

was quenched by adding water and then concentrated *in vacuo*. The resulting residue was dissolved in ethyl acetate. The organic solution was washed with water (2×) and brine (1×), and then dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation under reduced pressure, the crude product was purified using flash chromatography (silica gel; EtOAc only) to yield **2-29** ($C_{13}H_{28}O_6Si$, 2.80 g, 21%) as a white solid; TLC (EtOAc only) R_f = 0.22.

Subsequent benzylation of compound **2-29** (5.39 g, 17.5 mmol) using sodium hydride (60%, 2.52 g, 63.0 mmol) and benzyl bromide (9.0 mL, 75 mmol) in anhydrous DMF (50 mL) yielded **2-30** ($C_{34}H_{46}O_6Si$, 6.07 g, 60%) as a white solid; TLC (EtOAc/petroleum ether (1:9)) R_f = 0.39.

Removal of the silyl protecting group at C-6 of compound **2-30** (0.29 g, 0.50 mmol) was achieved using a tetrabutylammonium fluoride solution (1.0 M in THF, 1.5 mmol) in anhydrous THF (200 mL). The crude product was purified using flash chromatography (silica gel; EtOAc/hexanes (1:1)) to give **2-31** ($C_{28}H_{32}O_6$; 0.16 g, 69%) as a colorless solid; TLC (EtOAc/hexanes (1:1)) R_f = 0.32. ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) 7.42–7.24 (15 H, m, Bn×3), 4.97 (1 H, d, J = 11.6 Hz, CH_2 of Bn), 4.89 (1 H, d, J = 12.0 Hz, CH_2 of Bn), 4.84 (1 H, d, J = 12.0 Hz, CH_2 of Bn), 4.75 (1 H, d, J = 11.6 Hz, CH_2 of Bn), 4.70 (1 H, d, $J_{1,2}$ = 4.0 Hz, H1), 4.69 (1 H, d, J = 11.6 Hz, CH_2 of Bn), 4.64 (1 H, d, J = 11.6 Hz, CH_2 of Bn), 4.05 (1 H, dd, $J_{1,2}$ = 4.0 Hz, $J_{2,3}$ = 10 Hz, H2), 3.94 (1 H, dd, $J_{2,3}$ = 10 Hz, $J_{3,4}$ = 2.8 Hz, H3), 3.87 (1 H, d, $J_{3,4}$ = $J_{4,5}$ = 2.8 Hz, H4), 3.743.68 (2 H, m, H6), 3.52–3.44 (1 H, m, H5), 3.36 (3 H, s, OMe). The ^{1}H NMR spectral data of **2-31** is consistent with those previously reported.[121]

Methyl 2,3,4-tri-*O*-benzyl-α-D-*galacto*-hexodialdo-1,5-pyranoside (2-32).[121]

Following a similar procedure used for the preparation of the aldehyde **2-20**, compound **2-31** (2.60 g, 5.60 mmol) was converted to **2-32** by reacting with oxalyl chloride (1.0 mL, 11.46 mmol) and triethylamine (3.9 mL, 28.0 mmol) in anhydrous dichloromethane (10 mL) and DMSO (1.7 mL, 21.98 mmol). The crude aldehyde

2-32 was used for later steps without further purification.

Methyl 2,3,4-tri-*O*-benzyl-7,8-dideoxy-D-*glycero*-α-D-*galacto*-oct-7-enopyrano -side (2-33).

Following a similar procedure used for the preparation of **2-21**, aldehyde **2-32** ($\sim$ 5.60 mmol) generated as described above was reacted with vinylmagnesium bromide (0.7 M in THF, 30 mL) in anhydrous THF (50 mL). The reaction product was purified by flash chromatography (silica gel) using EtOAc/hexanes (1:4$\rightarrow$1:2) to elute the desired product **2-33** ($C_{30}H_{34}O_6$; 0.21 g, 8%) as a white solid; TLC (EtOAc/hexanes (1:1)) R_f = 0.54. The epimeric L-*glycero*-α-D-*galacto*-isomer **2-34** (0.41 g, 15%) was eluted using EtOAc/hexanes (1:1). ^{1}H NMR ($CDCl_3$, 400 MHz) of **2-33** δ (ppm) 7.40–7.24 (15 H, m, Bn$\times$3), 5.91 (1 H, ddd, $J_{6,7}$ = 5.2 Hz, $J_{7,8cis}$ = 10.4 Hz, $J_{7,8trans}$ = 17.2 Hz, H7), 5.32 (1 H, ddd, $J_{6,8trans}$ = $J_{8cis,8trans}$ = 1.6 Hz, $J_{7,8trans}$ = 17.2 Hz, $H8_{trans}$), 5.17 (1 H, ddd, $J_{6,8cis}$ = $J_{8cis,8trans}$ = 1.6 Hz, $J_{7,8cis}$ = 10.4 Hz, $H8_{cis}$), 5.02 (1 H, d, J = 11.6 Hz, CH_2 of Bn), 4.86 (1 H, d, J = 11.6 Hz, CH_2 of Bn), 4.82 (1 H, d, J = 12.4 Hz, CH_2 of Bn), 4.75 (1 H, d, J = 11.6 Hz, CH_2 of Bn), 4.68 (1 H, d, J = 12.4 Hz, CH_2 of Bn), 4.68 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.65 (1 H, d, J = 11.6 Hz, CH_2 of Bn), 4.22 (1 H, m, H6), 4.13 (1 H, dd, $J_{4,5}$ = 1.2, $J_{3,4}$ = 2.8 Hz, H4), 4.05 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{2,3}$ = 10 Hz, H2), 3.90 (1 H, dd, $J_{3,4}$ = 2.8 Hz, $J_{2,3}$ = 10 Hz, H3), 3.44 (1 H, dd, $J_{4,5}$ = 1.2 Hz, $J_{5,6}$ = 7.6 Hz, H5), 3.31 (3 H, s, OMe). ^{13}C NMR ($CDCl_3$, 100 MHz) δ (ppm) 138.6, 138.3, 138.0 128.4, 128.33, 128.26, 128.1, 127.9, 127.7, 127.5, 127.4, 115.6, 98.8, 79.2, 76.1, 74.39, 74.37, 73.5, 73.3, 72.1, 70.7, 55.2. ESI-HRMS Calcd. for $C_{30}H_{34}O_6[M + Na]^+$ 513.2253, Found: 513.2247. The configuration at C-6 position of **2-33** was determined by Mosher's ester analysis as described in Section 2.2.6.

The undesired epimer L-*glycero*-α-D-*galacto*-isomer **2-34**: 7.39–7.24 (15 H, m, Bn$\times$3), 5.68 (1 H, ddd, $J_{6,7}$ = 6.8 Hz, $J_{7,8cis}$ = 10.4 Hz, $J_{7,8trans}$ = 17.2 Hz, H7), 5.30 (1 H, ddd, $J_{6,8trans}$ = $J_{8cis,8trans}$ = 1.6 Hz, $J_{7,8trans}$ = 17.2 Hz, $H8_{trans}$), 5.18 (1 H, ddd, $J_{6,8cis}$ = $J_{8cis,8trans}$ = 1.6 Hz, $J_{7,8cis}$ = 10.4 Hz, $H8_{cis}$), 5.08 (1 H, d, J = 10.8 Hz, CH_2 of Bn), 4.88

(1 H, d, J = 11.6 Hz, CH_2 of Bn), 4.82 (1 H, d, J = 12.4 Hz, CH_2 of Bn), 4.75 (1 H, d, J = 11.6 Hz, CH_2 of Bn), 4.73 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.68 (1 H, d, J = 12.4 Hz, CH_2 of Bn), 4.57 (1 H, d, J = 10.8 Hz, CH_2 of Bn), 4.35 (1 H, m, H6), 4.06 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{2,3}$ = 9.6 Hz, H2), 3.95 (1 H, dd, $J_{4,5}$ = 1.2 Hz, $J_{3,4}$ = 2.8 Hz, H4), 3.92 (1 H, dd, $J_{3,4}$ = 2.8 Hz, $J_{2,3}$ = 9.6 Hz, H3), 3.51 (1 H, m, H5), 3.36 (3 H, s, OMe).

Methyl 6-*O*-2,3,4-tri-*O*-benzyl-7,8-dideoxy-D-*glycero*-α-D-*galacto*-oct-7-enopyra-noside (2-35).

Alcohol **2-33** (0.21 g, 0.43 mmol) was dissolved in anhydrous pyridine (3 mL) and added acetic anhydride (3 mL). The reaction mixture was stirred at room temperature for 16 h. The reaction was quenched by adding ethanol slowly and then concentrated in vacuo. The crude product was purified by flash chromatography (silica gel; EtOAc/hexanes (1:4)) to yield **2-35** ($C_{32}H_{36}O_7$, 0.17 g, 75%) as a light yellow syrup. ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) 7.42–7.26 (15 H, m, Bn×3), 5.88 (1 H, ddd, $J_{6,7}$ = 6.0 Hz, $J_{7,8cis}$ =10.4 Hz, $J_{7,8trans}$ = 17.2 Hz, H7), 5.42–5.38 (1 H, m, H6), 5.27 (1 H, ddd, $J_{6,8trans}$ = $J_{8cis,8trans}$ = 1.2 Hz, $J_{7,8trans}$ = 17.2 Hz, $H8_{trans}$), 5.22 (1 H, ddd, $J_{6,8cis}$ = $J_{8cis,8trans}$ = 1.2 Hz, $J_{7,8cis}$ = 10.4 Hz, $H8_{cis}$), 4.94 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.91 (1 H, J = 12.0 Hz, CH_2 of Bn), 4.82 (1 H, J = 12.4 Hz, CH_2 of Bn), 4.77 (1 H, J = 12.0 Hz, CH_2 of Bn), 4.68 (1 H, J = 12.4 Hz, CH_2 of Bn), 4.66 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.51 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.06 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{2,3}$ = 9.6 Hz, H2), 3.96–3.93 (2 H, m, H3, H4), 3.67 (1 H, d, $J_{5,6}$ = 9.2, H5), 3.31 (3 H, s, OMe), 1.98 (3 H, s, OAc). ^{13}C NMR ($CDCl_3$, 100 MHz) δ (ppm) 169.0, 138.7, 138.3, 138.1, 134.2, 128.52, 128.33, 128.30, 128.2, 128.1, 127.7, 127.6, 127.5, 127.4, 117.6, 98.7, 79.4, 76.2, 74.6, 74.0, 73.54, 73.51, 71.4, 71.3, 55.2, 21.0. ESI-HRMS Calcd. for $C_{32}H_{36}O_7[M + Na]^+$ 555.2359, Found: 555.2351.

Methyl 6-*O*-acetyl-2,3,4-tri-*O*-benzyl-α-D-*erythro*-D-*galacto*-octopyranoside (2-36).

The same procedure used to prepare **2-24** was followed to convert compound **2-35**

(0.17 g, 0.32 mmol) to **2-36** using AD-mix α (0.50 g) and methanesulfonamide (30 mg, 0.32 mmol) in 5 mL of co-solvent ^{t}BuOH : H_2O (1:1). The crude diol **2-36** was used for the next step without further purification.

Methyl 6,7,8-tri-*O*-acetyl-2,3,4-tri-*O*-benzyl-α-D-*erythro*-D-*galacto*-octopyrano -side (2-37).

The above crude diol **2-36** (0.32 mmol) was reacted with acetic anhyide (5 mL) in anhydrous pyridine (5 mL). The crude **2-37** was used for the next step without further purification.

1,2,3,4,6,7,8-Hepta-*O*-acetyl-α,β-D-*erythro*-D-*galacto*-octopyranose (2-39).

Similar to the method used to prepare compound **2-27**, compound **2-37** and 5% Pd/C (50 mg) in a methanol solution (5 mL) were stirred under hydrogen atmosphere (1 atm) for 16 h. The reaction mixture was filtered through Celite and the filtrate after concentration was reacted with acetic anhydride (5 mL) in anhydrous pyridine (5 mL). The resulting intermediate **2-38** (0.14 g, 0.22 mmol) was subjected to acetic acid (3 mL), acetic anhydride (3 mL) and concentrated sulfuric acid (0.4 mL). After stirring at 4 °C in cold room for 16 h, the reaction was quenched by adding saturated sodium bicarbonate solution. After routine work-up, purification of the organic extracts using flash chromatography (silica gel; EtOAc/hexanes (1:1)) gave **2-39** ($C_{22}H_{30}O_{15}$; 0.075 g, 71%) in a white solid form as a mixture of α : β isomers (9:2 ratio). ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) α anomer: 6.35 (1 H, d, $J_{1,2}$ = 3.0 Hz, H1), 5.43 (1 H, dd, $J_{6,7}$ = 3.0 Hz, $J_{5,6}$ = 1.5 Hz, H6), 5.32–5.27 (3 H, m, H2, H3 and H7), 5.18 (1 H, dd, $J_{4,5}$ = 10.0 Hz, $J_{3,4}$ = 3.0 Hz, H4), 4.33 (1 H, dd, $J_{4,5}$ = 10.0 Hz, $J_{5,6}$ = 1.5 Hz, H5), 4.27 (1 H, dd, $J_{7,8a}$ = 3.0 Hz, $J_{8a,8b}$ = 12.0 Hz, H_{8a}), 4.10 (1 H, dd, $J_{7,8b}$ = 8.0 Hz, $J_{8a,8b}$ = 12.0 Hz, H_{8b}), 2.17 (3 H, s, OAc), 2.08 (3 H, s, OAc), 2.04 (3 H, s, OAc), 2.02 (3 H, s, OAc), 2.01 (3 H, s, OAc), 1.97 (3 H, s, OAc), 1.96 (3 H, s, OAc). ^{13}C NMR ($CDCl_3$, 100 MHz) δ (ppm) α anomer: 170.7, 170.6, 170.4, 170.04, 169.99, 169.4, 169.0, 89.6, 71.0, 70.0, 67.9, 67.5, 66.6, 66.4, 62.0,

21.00, 20.94, 20.79, 20.77, 20.73, 20.67. ESI-HRMS Calcd. for $C_{22}H_{30}O_{15}[M + Na]^+$ 557.1482; Found: 557.1477.

^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) β anomer: 5.62 (1 H, d, $J_{1,2}$ = 8.0 Hz, H1), 5.38 (1 H, dd, $J_{6,7}$ = 3.5 Hz, $J_{5,6}$ = 1.5 Hz, H6), 5.32–5.27 (3 H, m, H2, H3 and H7), 5.06 (1 H, dd, $J_{3,4}$ = 3.5 Hz, $J_{4,5}$ = 10.5 Hz, H4), 4.34–4.32 (1 H, m, H5), 4.16 (1 H, dd, $J_{7,8a}$ = 5.5 Hz, $J_{8a,8b}$ = 12.0 Hz, H_{8a}), 3.91 (1 H, dd, $J_{7,8b}$ = 7.5 Hz, $J_{8a,8b}$ = 12.0 Hz, H_{8b}), 2.12 (3 H, s, OAc), 2.10 (3 H, s, OAc), 2.07 (3 H, s, OAc), 2.03 (3 H, s, OAc), 2.03 (3 H, s, OAc), 2.00 (3 H, s, OAc), 1.95 (3 H, s, OAc).

2.2.6 Mosher's Ester Analysis

Determination of C6-(*R*) configuration of compound 2-21.

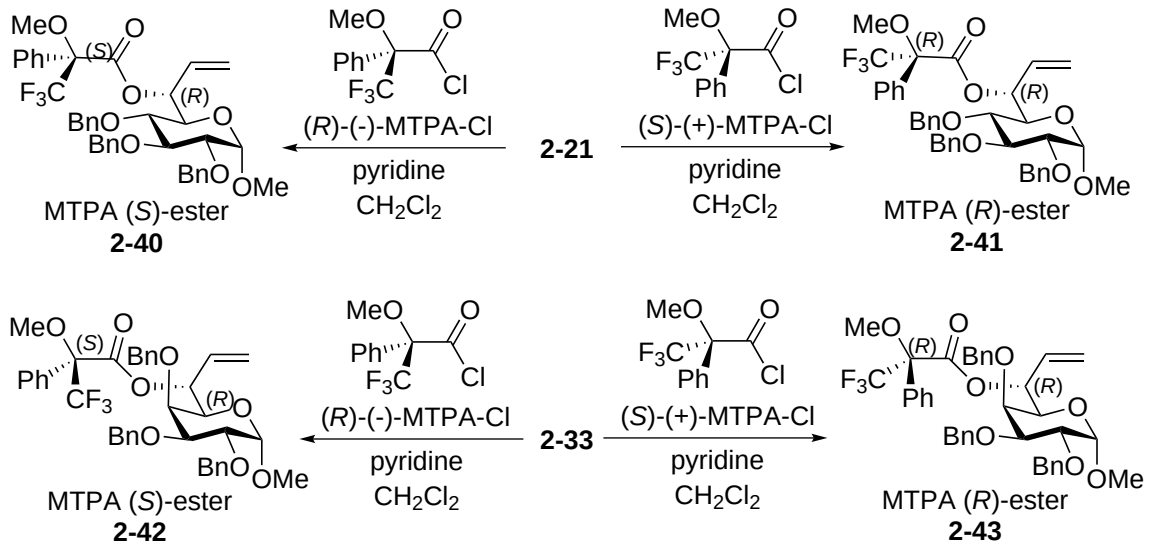


Figure 2.5: Preparation of (*S*)- and (*R*)-MTPA esters of **2-21** and **2-33** for stereochemical analysis.

To determine the C6 configuration of **2-21**, the (*S*)- and (*R*)-MTPA (α-methoxy-α-trifluoromethylphenylacetic acid) ester (**2-40** and **2-41**) were prepared (Figure 2.5, Table 2.1).[122] To a stirred solution of alcohol **2-21** (75.5 mg, 0.15 mmol) and anhydrous pyridine (0.4 mL) in anhydrous chloroform (1 mL) at room temperature, *R*-(−)-MTPA-Cl (61 mg, 0.24 mmol) was added. The reaction mixture was

Table 2.1: Mosher's ester analysis for C-6 configuration determination of **2-21** and **2-33**. $\Delta\delta = \delta_S - \delta_R$.

	δ_S **2-40** (ppm)	δ_R **2-41** (ppm)	$\Delta\delta$	δ_S **2-42** (ppm)	δ_R **2-43** (ppm)	$\Delta\delta$
H7	5.89	5.79	+ 0.10	5.82	5.78	+ 0.04
$H8_{trans}$	5.27	5.23	+ 0.04	5.41	5.30	+ 0.11
$H8_{cis}$	5.33	5.23	+ 0.10	5.34	5.26	+ 0.08
H1	4.50	4.56	− 0.06	4.58	4.60	− 0.02
H3	3.96	4.00	− 0.04	3.82	3.87	− 0.05
H5	3.77	3.91	− 0.14	3.70	3.73	− 0.03
OMe	3.01	3.29	− 0.28	3.25	3.27	− 0.02
H2	3.37	3.42	− 0.05	3.97	4.00	− 0.03
H4	3.28	3.39	− 0.11	3.70	3.79	− 0.09

stirred at room temperature and the progress of the reaction was monitored by TLC (EtOAc/hexanes (1:4)). After the consumption of the starting material was completed (2 h), the reaction was quenched by adding water and dichloromethane. The aqueous layer was extracted with dichloromethane (3×) and the combined organic layers were dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation, the crude product was purified using flash chromatography (silica gel; EtOAc/hexanes (1:9→1:4)) to yield the (*S*)-MTPA ester **2-40** ($C_{40}H_{41}F_3O_8$) as a colorless solid; TLC (EtOAc/hexanes (1:4)) R_f = 0.53. ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) 7.51–7.26 (20 H, m, Bn×3, Ph-Hs of ester), 5.89 (1 H, ddd, $J_{6,7}$ = 8.8 Hz, $J_{7,8cis}$ = 10.4 Hz, $J_{7,8trans}$ = 17.2 Hz, H7), 5.82 (1 H, ddd, $J_{6,8cis}$ = 1.6 Hz, $J_{5,6}$ = 2.0 Hz, $J_{6,7}$ = 8.8, H6), 5.33 (1 H, ddd, $J_{6,8cis}$ = $J_{8cis,8trans}$ = 1.6 Hz, $J_{7,8cis}$ =10.4 Hz, H_{8cis}), 5.27 (1 H, ddd, $J_{6,8trans}$ = $J_{8cis,8trans}$ = 1.6 Hz, $J_{7,8trans}$ = 17.2 Hz, H_{8trans}), 4.97 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.86 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.76 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.74 (1 H, J = 12.4 Hz, CH_2 of Bn), 4.61 (1 H, J = 12.4 Hz, CH_2 of Bn), 4.55 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.50 (1 H, d, $J_{1,2}$ = 3.2 Hz, H1), 3.96 (1 H, dd, $J_{3,4}$ = 8.8 Hz, $J_{2,3}$ = 9.6 Hz, H3), 3.77 (1 H, dd, $J_{5,6}$ = 2.0 Hz, $J_{4,5}$ = 10.0 Hz, H5), 3.53 (3 H, s, OMe of ester), 3.37 (1 H, dd, $J_{1,2}$ = 3.2 Hz, $J_{2,3}$ = 9.6 Hz,

H2), 3.28 (1 H, dd, $J_{3,4}$ = 8.8 Hz, $J_{4,5}$ = 10.0 Hz, H4), 3.01 (3 H, s, OMe).

In an analogous fashion, the (*R*)-MTPA ester **2-41** was prepared using (*S*)-(+)-MTPA-Cl. ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) 7.51–7.23 (20 H, m, Bn×3, Ph-Hs of ester), 5.79 (1 H, ddd, $J_{6,7}$ = 7.6 Hz, $J_{7,8cis}$ = 16.4 Hz, $J_{7,8trans}$ = 20.0 Hz, H7), 5.79 (1 H, m, H6), 5.23 (1 H, m, H_{8trans}), 5.23 (1 H, m, H_{8cis}), 4.98 (1 H, *J* = 10.8 Hz, CH_2 of Bn), 4.84 (1 H, *J* = 10.8 Hz, CH_2 of Bn), 4.78 (1 H, *J* = 10.8 Hz, CH_2 of Bn), 4.78 (1 H, *J* = 12.4 Hz, CH_2 of Bn), 4.64 (1 H, *J* = 12.4 Hz, CH_2 of Bn), 4.56 (1 H, d, $J_{1,2}$ = 3.2 Hz, H1), 4.51 (1 H, *J* = 10.8 Hz, CH_2 of Bn), 4.00 (1 H, dd, $J_{3,4}$ = 8.8 Hz, $J_{2,3}$ = 9.2 Hz, H3), 3.91 (1 H, dd, $J_{5,6}$ = 1.2 Hz, $J_{4,5}$ = 10.4 Hz, H5), 3.50 (3 H, s, OMe of ester), 3.42 (1 H, dd, $J_{1,2}$ = 3.2 Hz, $J_{2,3}$ = 9.2 Hz, H2), 3.39 (1 H, dd, $J_{3,4}$ = 8.8 Hz, $J_{4,5}$ = 10.4 Hz, H4), 3.29 (3 H, s, OMe).

Determination of C6-(*R*) configuration of compound 2-33.

The (*S*)- and (*R*)-MTPA esters (**2-42** and **2-43**) of **2-33** (63 mg, 0.13 mmol) were prepared according to the same procedure described above for the synthesis of **2-40** and **2-41** using *R*-(−)-MTPA-Cl (61 mg, 0.24 mmol) (Figure 2.5, Table 2.1).[122] The reaction mixture was purified using flash chromatography (silica gel; EtOAc/hexanes (1:9)) to yield (*S*)-MTPA ester **2-42** ($C_{40}H_{41}F_3O_8$) as a colorless solid; TLC (EtOAc/hexanes (1:4)) R_f = 0.30. ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) 7.52–7.25 (20 H, m, Bn×3, Ph-Hs of ester), 5.82 (1 H, ddd, $J_{6,7}$ = 7.6 Hz, $J_{7,8cis}$ = 10.4 Hz, $J_{7,8trans}$ = 17.2 Hz, H7), 5.66 (1 H, ddd, $J_{6,8cis}$ = 1.2 Hz, $J_{6,7}$ = 7.6 Hz, $J_{5,6}$ = 9.2 Hz, H6), 5.41 (1 H, ddd, $J_{6,8trans} = J_{8cis,8trans}$ = 1.2 Hz, $J_{7,8trans}$ =17.2 Hz, H_{8trans}), 5.34 (1 H, ddd, $J_{6,8cis} = J_{8cis,8trans}$ = 1.2 Hz, $J_{7,8cis}$ = 10.4 Hz, H_{8cis}), 4.89 (1 H, *J* = 10.8 Hz, CH_2 of Bn), 4.78 (1 H, *J* = 12.4 Hz, CH_2 of Bn), 4.74 (1 H, *J* = 12.0 Hz, CH_2 of Bn), 4.64 (1 H, *J* = 12.0 Hz, CH_2 of Bn), 4.64 (1 H, *J* = 12.4 Hz, CH_2 of Bn), 4.58 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.27 (1 H, *J* = 10.8 Hz, CH_2 of Bn), 3.97 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{2,3}$ = 10.4 Hz, H2), 3.82 (1 H, dd, $J_{3,4}$ = 2.8 Hz, $J_{2,3}$ = 10.4 Hz, H3), 3.70–3.69 (2 H, m,

H4, H5), 3.50 (3 H, s, OMe of ester), 3.25 (3 H, s, OMe).

The (*R*)-MTPA ester **2-43** was prepared in a same manner using (*S*)-(+)-MTPA-Cl. ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) 7.48–7.25 (20 H, m, Bn×3, Ph-Hs of ester), 5.78 (1 H, ddd, $J_{6,7}$ = 7.6 Hz, $J_{7,8cis}$ = 10.4 Hz, $J_{7,8trans}$ = 17.6 Hz, H7), 5.63 (1 H, ddd, $J_{6,8cis}$ = 1.2 Hz, $J_{6,7}$ = 7.6 Hz, $J_{5,6}$ = 9.2 Hz, H6), 5.30 (1 H, ddd, $J_{6,8trans}$ = $J_{8cis,8trans}$ = 1.2 Hz, $J_{7,8trans}$ =17.6 Hz, H_{8trans}), 5.26 (1 H, ddd, $J_{6,8cis}$ = $J_{8cis,8trans}$ = 1.2 Hz, $J_{7,8cis}$ = 10.4 Hz, H_{8cis}), 4.92 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.82 (1 H, J = 12.0 Hz, CH_2 of Bn), 4.79 (1 H, J = 12.8 Hz, CH_2 of Bn), 4.68 (1 H, J = 12.8 Hz, CH_2 of Bn), 4.65 (1 H, J = 12.0 Hz, CH_2 of Bn), 4.60 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.32 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.00 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{2,3}$ = 10.0 Hz, H2), 3.87 (1 H, dd, $J_{3,4}$ = 2.4 Hz, $J_{2,3}$ = 10.0 Hz, H3), 3.79–3.78 (1 H, m, H4), 3.73 (1 H, broad d, $J_{5,6}$ = 9.2 Hz, H5), 3.42 (3 H, s, OMe of ester), 3.27 (3 H, s, OMe).

Determination of C7-(*R*) configuration of compound 2-24.

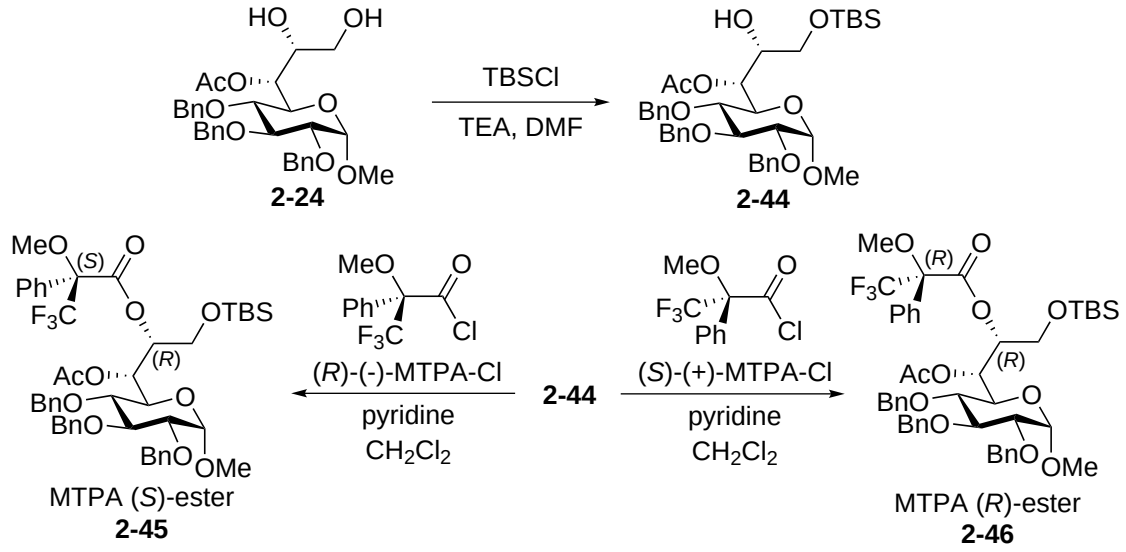


Figure 2.6: Derivatization of diols **2-24** and preparation of (*S*)- and (*R*)-MTPA esters of **2-44** for stereochemical analysis.

To determine the C7 configuration of **2-24**, this diol compound was first converted to **2-44** and then derivatized to make the (*S*)- and (*R*)-MTPA esters **2-**

Table 2.2: Mosher's ester analysis for C-7 configuration determination of **2-24**. $\Delta\delta = \delta_S - \delta_R$.

	δ_S **2-45** (ppm)	δ_R **2-46** (ppm)	$\Delta\delta$
OMe	3.21	3.35	−0.14
H1	4.53	4.57	−0.04
H2	3.22	3.35	−0.13
H3	3.83	3.91	−0.08
H4	3.1	3.42	−0.32
H5	3.98	4.11	−0.13
H6	5.48	5.40	0.08
OAc	2.02	2.00	0.02
H7	4.21	4.13	0.08
$H8_a$	4.23	4.02	0.21
$H8_b$	3.79	3.64	0.15
^{t}Bu of TBS	0.81	0.78	0.03
Me_1 of TBS	−0.093	−0.18	0.087
Me_2 of TBS	−0.193	−0.34	0.147

45 and **2-46** (Figure 2.6, Table 2.2).[122] Specifically, compound **2-24** (230 mg, 0.41 mmol) was dissolved in anhydrous DMF (10 mL) and the reaction mixture was cooled to 0 °C. To this solution was added triethylamine (0.20 mL, 1.4 mmol), and the resulting mixture was stirred at 0 °C for 10 min. TBSCl (0.080 g, 0.53 mmol) was added and the reaction mixture was brought to room temperature slowly. After stirring for 16 h, the reaction was quenched by adding water and concentrated in vacuo. The residue was dissolved in ethyl acetate and the organic solution was washed with water (1×) and brine (1×), and then dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation under reduced pressure, the crude product was purified using flash chromatography (silica gel; EtOAc/hexanes (1:4→1:2)) to yield **2-44** ($C_{38}H_{52}O_9Si$; 0.11 g, 40%) as a colorless syrup. ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) 7.48–7.37 (15 H, m, Bn×), 5.13 (1 H, J = 10.8 Hz, CH_2 of Bn), 5.10 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.95 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.90 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.80 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.77 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.67 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.32–4.20 (2 H,

m, H_{8a}, H_{8b}), 4.16 (1 H, ddd, $J_{7,8a}$ = 2.0 Hz, $J_{6,7}$ = 6.0 Hz, $J_{7,8b}$ = 8.0 Hz, H7), 4.22 (1 H, dd, $J_{2,3}$ = $J_{3,4}$ = 9.2 Hz, H3), 4.06 (1 H, dd, $J_{4,5}$ = 9.2 Hz, $J_{5,6}$ = 4.0 Hz, H5), 3.82 (1 H, dd, $J_{4,5}$ = $J_{3,4}$ = 9.2 Hz, H4), 3.75 (1 H, dd, $J_{6,7}$ = 6.0 Hz, $J_{5,6}$ = 4.0 Hz, H6), 3.63 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{2,3}$ = 9.2 Hz, H2), 3.51 (3 H, s, OMe), 2.16 (3 H, s, OAc), 0.97 (9 H, s, ^{t}Bu of TBS), 0.095 (3 H, s, Me of TBS), 0.022 (3 H, s, Me of TBS). ^{13}C NMR ($CDCl_3$, 125 MHz) δ (ppm) 171.3, 171.1, 138.6, 137.97, 137.8, 128.4, 128.3, 128.04, 128.01, 127.9, 127.8, 127.7, 127.5, 97.9, 82.3, 79.9, 79.6, 75.6, 74.8, 74.5, 73.2, 71.9, 67.8, 66.4, 60.3, 55.5, 25.7, 21.0, 20.9, 18.0, 14.1, 4.7, 5.0. ESI-HRMS Calcd. for $C_{38}H_{52}O_9Si[M + Na]^+$ 703.3278, Found: 703.3276.

The (*S*)- and (*R*)-MTPA esters (**2-45** and **2-46**) of **44** (63 mg, 0.13 mmol) were prepared according to the same procedure described above for the synthesis of **2-40** and **2-41** using *R*-(−)-MTPA-Cl (61 mg, 0.24 mmol), respectively. ^{1}H NMR ($CDCl_3$, 400 MHz) of **2-45**: δ (ppm) 7.57–7.18 (20 H, m, Bn×3, Ph-Hs of ester), 5.48 (1 H, broad d, $J_{6,7}$ = 8.0 Hz, H6), 4.94 (1 H, *J* = 11.0 Hz, CH_2 of Bn), 4.75 (1 H, *J* = 11.0 Hz, CH_2 of Bn), 4.75 (1 H, *J* = 10.5 Hz, CH_2 of Bn), 4.71 (1 H, *J* = 12.0 Hz, CH_2 of Bn), 4.63 (1 H, *J* = 12.0 Hz, CH_2 of Bn), 4.53 (1 H, d, $J_{1,2}$ = 3.5 Hz, H1), 4.33 (1 H, *J* = 10.5 Hz, CH_2 of Bn), 4.23 (1 H, dd, $J_{7,8a}$ = 2.0 Hz, $J_{8a,8b}$ = 12.0 Hz, H_{8a}), 4.21 (1 H, ddd, $J_{7,8a}$ = 2.0 Hz, $J_{7,8b}$ = 5.5 Hz, $J_{6,7}$ = 8.0 Hz, H7), 3.98 (1 H, broad d, $J_{4,5}$ = 10.0 Hz, H5), 3.83 (1 H, dd, $J_{3,4}$ = 9.5 Hz, $J_{2,3}$ = 10.5 Hz, H3), 3.79 (1 H, dd, $J_{7,8b}$ = 5.5 Hz, $J_{8a,8b}$ = 12.0 Hz, H_{8b}), 3.62 (3 H, s, OMe of ester), 3.22 (1 H, dd, $J_{1,2}$ = 3.5 Hz, $J_{2,3}$ = 10.5 Hz, H2), 3.10 (1 H, dd, $J_{3,4}$ = 9.5 Hz, $J_{4,5}$ = 10.0 Hz, H4), 3.21 (3 H, s, OMe), 2.02 (3 H, s, OAc), 0.81 (9 H, s, ^{t}Bu of TBS), −0.093 (3 H, s, Me of TBS), −0.193 (3 H, s, Me of TBS).

^{1}H NMR ($CDCl_3$, 400 MHz) of **2-46**: δ (ppm) 7.53–6.98 (20 H, m, Bn×3, Ph-Hs of ester), 5.40 (1 H, broad d, $J_{6,7}$ = 9.2 Hz, H6), 4.96 (1 H, *J* = 10.8 Hz, CH_2 of Bn), 4.81 (1 H, *J* = 10.8 Hz, CH_2 of Bn), 4.79 (1 H, *J* = 10.4 Hz, CH_2 of Bn), 4.77 (1 H, *J* = 12.4 Hz, CH_2 of Bn), 4.66 (1 H, *J* = 12.4 Hz, CH_2 of Bn), 4.57 (1 H, d, $J_{1,2}$ = 3.2 Hz, H1),

4.24 (1 H, J = 10.4 Hz, CH_2 of Bn), 4.13 (1 H, ddd, $J_{7,8a}$ = 2.4 Hz, $J_{7,8b}$ = 4.8 Hz, $J_{6,7}$ = 9.2 Hz, H7), 4.11 (1 H, broad d, $J_{4,5}$ = 9.6 Hz, H5), 4.02 (1 H, dd, $J_{7,8a}$ = 2.4 Hz, $J_{8a,8b}$ = 12.0 Hz, H_{8a}), 3.91 (1 H, dd, $J_{3,4}$ = 9.6 Hz, $J_{2,3}$ = 9.6 Hz, H3), 3.64 (1 H, dd, $J_{7,8b}$ = 4.8 Hz, $J_{8a,8b}$ = 12.0 Hz, H_{8b}), 3.50 (3 H, s, OMe of ester), 3.42 (1 H, dd, $J_{3,4}$ = 9.6 Hz, $J_{4,5}$ = 9.6 Hz, H4), 3.35 (1 H, dd, $J_{1,2}$ = 3.2 Hz, $J_{2,3}$ = 9.6 Hz, H2), 3.35 (3 H, s, OMe), 2.00 (3 H, s, OAc), 0.78 (9 H, s, tBu of TBS), -0.18 (3 H, s, Me of TBS), -0.34 (3 H, s, Me of TBS).

2.2.7 Peracetylation of the LmbN Products

The purified *C*-His_6-LmbR and *C*-His_6-LmbN were dialyzed against 100 mM NH_4HCO_3 buffer containing 10 mM $MgCl_2$ at 4 °C to remove glycerol and salts. To the enzyme solution was then added R5P (**2-52**, 10 mM) and F6P (**2-12**, 10 mM). The final concentration of *C*-His_6-LmbR and *C*-His_6-LmbN were 40 μM and 50 μM, respectively. As a control experiment, the reaction mixture containing only *C*-His_6-LmbR and the sugar substrates was similarly prepared. Each reaction mixture (400 μL) was incubated at 30 °C for 4.5 h and an additional 3.5 h at 37 °C. The mixture was then added calf intestinal alkaline phosphatase (CIP)(4 μL, 40 units), and the reaction was incubated at 37 °C for 12 h. The resulting solution was added MeOH (400 μL) and MeCN (800 μL). After mixing the solution, the precitipates were removed by centrifugation at 16,000 $\times$ g for 5 min. The supernatant (400 μL) was evaporated *in vacuo* using a Speedvac SC100 (Savant). The resultant residue was completely dried and then treated with anhydrous pyridine (20 μL) and acetic anhydride (20 μL). This mixture was incubated at room temperature for 12 h, and quenched by the addition of water (150 μL). The peracetylated products were extracted with ethyl acetate (150 μL) and subjected to HPLC analysis using an analytical C_{18} column (4×250 mm). The sample (5 μL) was eluted with a gradient of water (solvent A) and 80% MeCN (solvent B). The gradient was run from

40-60% B over 15 min, 60-100% B over 5 min with a 3 min wash at 100% B, 100-40% B over 2 min, followed by re-equilibration at 40% B for 5 min. The flow rate was set at 1 mL/min and a corona CAD detector was used for the detection. Authentic standards of peracetylated octose derivatives were synthesized (Sections 2.2.4 and **??**) and analyzed using the same HPLC method. The recombinant LmbR and LmbN were prepared by another graduate student, Eita Sasaki.

2.3 Results and Discussion

2.3.1 Proposed Biosynthetic Pathways of Lincomycin A and Celesticetin

To date, two biosynthetic gene clusters from *S. lincolnensis* strains 78-11[115] and ATCC 25466[117] were identified to be responsible for the production of lincomycins. The overall organizations of these two gene clusters is very similar. The biosynthetic gene cluster of structurally related celesticetin from *S. caeletis* ATCC 1584 was disclosed using hybridization analysis.[118] The complete sequence of both clusters are available in the NCBI database (the accession numbers of lincomycin A and celesticetin biosynthetic gene clusters are EU124663 and GQ844764, respectively).

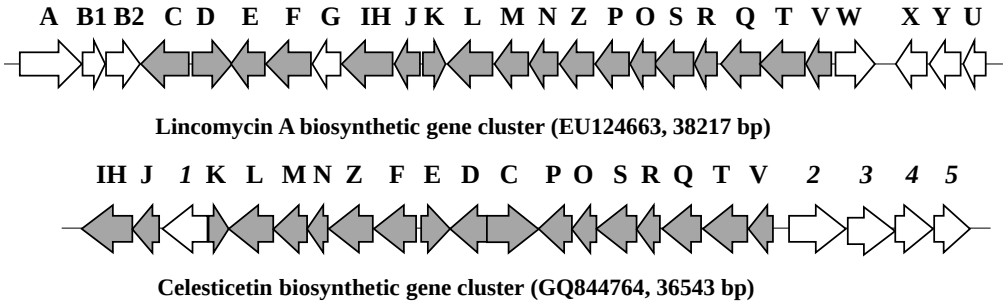


Figure 2.7: Biosynthetic gene clusters of lincomycin A and celesticetin. Homologous genes found in both cluster are shown in gray.

Table 2.3: Sequence homology of the proteins found in lincomycin A and celesticetin biosynthetic pathways

Lincomycin A	**Celesticetin**	**identity / similarity** (%)
LmbC	CcbC	56 / 66
LmbD	CcbD	56 / 65
LmbE	CcbE	60 / 71
LmbF	CcbF	40 / 55
LmbIH	CcbIH	63 / 71
LmbJ	CcbJ	62 / 71
LmbK	CcbK	64 / 72
LmbL	CcbL	64 / 75
LmbM	CcbM	74 / 82
LmbN	CcbN	72 / 80
LmbZ	CcbZ	64 / 75
LmbP	CcbP	59 / 70
LmbO	CcbO	65 / 74
LmbS	CcbS	74 / 81
LmbR	CcbR	69 / 80
LmbQ	CcbQ	44 / 54
LmbT	CcbT	65 / 76
LmbV	CcbV	56 / 68

Sequence analysis of the lincomycin and celesticetin biosynthetic gene clusters

Since the structures of lincomycin A and celesticetin share the same thiooctose core, the genes encoding enzymes involved in the biosynthesis of the thiosugar moiety should exist in both clusters and serve similar roles. In fact, sequence analysis has shown that there are eighteen homologous genes between *lmb* and *ccb* clusters (Figure 2.7, Table 2.3). Naturally, the structural differences of lincomycin A and celesticetin are reflected by the composition and organization of their respective biosynthetic gene clusters. The genes found exclusively in *lmb* cluster but not in *ccb* cluster are likely involved in the formation of the *N*-methyl-PPL (**2-2**) moiety and the *S*-methyltransfer at C-1 position. On the other hand, the genes which are absent in the lincomycin biosynthetic gene cluster are possibly associated with construction of the salicyloyloxyethyl (SaE, **2-6**) substituent attached to

the C-1 thiol group and the C-7 *O*-methylation of the octose (Figure 2.1). Here, the BLAST seach of the protein database was carried out using the primary amino acid sequence of each protein encoded in the lincomycin A and celesticetin biosynthetic gene clusters as queries. The results of the sequence similarity analysis are shown in Table 2.4 [†] and 2.5.

Table 2.4: Sequence analysis of proteins encoded in *lmb* gene cluster

Protein	**Entry number**	**Protein homolog**	**Entry number**	**identity / similarity** (%)
LmbA	A9Y8R0	Por11	W5QK70	73 / 80
LmbB1	Q54354	SibV	C0LTN6	57 / 65
LmbB2	A9Y8R3	D-Tyosine 3-hydroxylase	H2JL10	50 / 58
LmbC	A9Y8R4	Acyl-CoA synthetase	H2JPE3	41 / 52
LmbD	A9Y8R5	D-Lactate dehydrogenase	A0A085HJV1	37 / 52
LmbE	A9Y8R6	MshB	R4Z6V6	40 / 51
LmbF	A9Y8R7	Aminotransferase class I and II	D3Q619	29 / 43
LmbG	A9Y8R8	Methyltransferase	M7A8U6	33 / 44
LmbIH	Q9L6I6	Peptidase U62	Q1IK33	36 / 50
LmbJ	A9Y8S0	Methyltransferase type 11	D3D5Y3	42 / 52
LmbK	Q54364	Heptose 1,7-bisphosphate phosphatase	V4YJX7	51 / 63
LmbL	A9Y8S2	UDP-glucose 6-dehydrogenase	Q1MNW7	35 / 50
LmbM	A9Y8S3	NDP-sugar epimerase	H8I9P6	46 / 61
LmbN	A9Y8S4	Sedoheptulose 7-phosphate isomerase	A9GWB8	41 / 56
LmbZ	A9Y8S5	Oxidoreductase family	U2QFS7	44 / 59
LmbP	A9Y8S6	GHMP kinase	R4Z470	37 / 52
LmbO	A9Y8S7	Nucleotidyl transferase	W9HCS3	39 / 53
LmbS	A9Y8S8	PLP-dependent transferase	A0A0A1VLN6	49 / 62
LmbR	A9Y8S9	Transaldolase	D5UDL8	45 / 62
LmbQ	A9Y8T0	Protein pmbA	G8S2W9	29 / 40
LmbT	A9Y8T1	Glycosyltransferase	C4U8A9	24 / 43
LmbV	A9Y8T2	Uncharacterized protein	W2ETW6	43 / 51
LmbW	A9Y8T3	Methyltransferase	A8W7D8	78 / 85
LmbX	Q54378	TomK	C0LTT2	49 / 60
LmbY	Q54379	F-420 dependent reductase	H0BQJ4	55 / 67
LmbU	Q54380	Uncharacterized protein	A0A075V0A8	56 / 70

[†]The homologous proteins were searched in organisms excluding strains of *S. lincolnenses* and *S. caelestic*.

Table 2.5: Sequence analysis of proteins encoded in *ccb* gene cluster

Protein	**Entry number**	**Protein homolog**	**Entry number**	**identity / similarity** (%)
Ccb1	E9JES1	Diacylglycerol *O*-acyltransferase	W5QK70	35 / 44
Ccb2	E9JET8	CoA ligase	B0BLM7	56 / 67
Ccb3	E9JET9	Anthranilate synthase component I	A4FD46	57 / 69
Ccb4	E9JEU0	*O*-methyltransferase	D2QXN7	39 / 60
Ccb5	E9JEU1	Alcohol dehydrogenase	A0A024YV11	68 / 78

Biosynthesis of the 4-propyl-L-proline moiety (2-4)

Feeding experiments using L-[1-^{14}C]-tyrosine indicated that seven of the nine carbons in the propylproline unit are derived from tyrosine.[123] The other two carbons, including the *N*-methyl and terminal aliphatic side chain, are possibly transferred from SAM. The results of the radio-tracing experiments with D-[$^{13}C_6$]-glucose had led to a proposed pathway starting from L-tyrosine (**2-47**). The early steps may involve the formation of L-3,4-dihydroxyphenyl-alanine (DOPA, **2-48**), followed by a 2,3-extradiol cleavage to yield five-membered ring amino acid precursor of lincomycin A (**2-1**).[123] These two steps are likely catalyzed by LmbB2 and LmbB1. The availability of lincomycin A biosynthetic gene cluster allows the biochemical characterization of LmbB2 and LmbB1 (Table 2.4).[124,125]

These first two biosynthetic steps resemble those observed in the biosynthesis of pyrrolo[1,4]benzodiazepine antibiotics, sibiromycin, anthramycin, tomaymycin, and a highly modified depsipeptide hormaomycin (Table 2.6).[126–129] Comparison of their gene clusters led to the functional assignments of LmbX, LmbW, LmbA and LmbY (Figure 2.8). Specifically, LmbX may act as a potential hydrolase catalyzing the C–C bond cleavage of the LmbB1-product (**2-49**) to yield the diene (**2-50**), which is subsequently methylated by a putative SAM-dependent methyltrasnferase, LmbW. A mutant of the lincomycin producing strain impaired in the biosynthesis of F-420 cofactor accumulated the diene (**2-51**), suggesting that LmbY

(a putative F-420-dependent oxidoreductase) may mediate the reduction of the diene system.[130] LmbA, which is a putative γ-glutamyl transpeptidase, is possibly involved in the maturation process of F-420 cofactor.

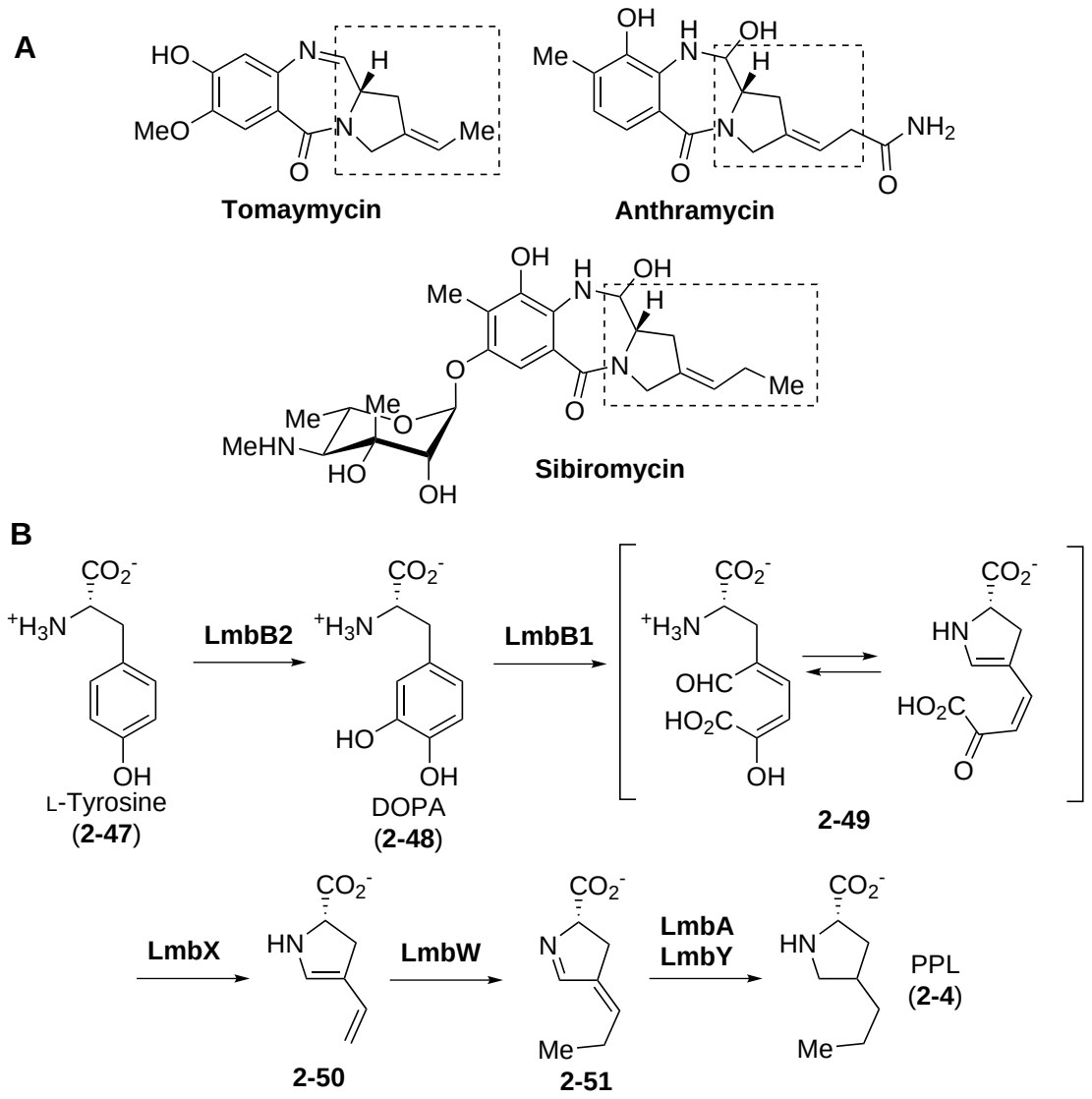


Figure 2.8: (A) Natural products containing the PPL moiety. (B) Proposed biosynthesis of the PPL moiety.

Biosynthetic precursors of the octose

The proposed overall biosynthetic pathway of lincomycin A is depicted in Figure 2.9. A putative transaldolase, LmbR, has been proposed to catalyze the con-

Table 2.6: Enzyme homologs from the lincomycin A, antramycin, sibiromycin, tomaymycin and hormaomycin biosynthetic pathways.

	Homologs found in the gene clusters					
Lincomycin A	LmbB2	LmbB1	LmbX	LmbW	LmbA	LmbY
Antramycin	ORF13	ORF12	ORF15	ORF5	ORF6	ORF14
Sibiromycin	SibU	SibV	SibS	SibZ	SibY	SibT
Tomaymycin	TomI	TomH	TomK		TomL	TomJ
Hormaomycin	HrmE	HrmF		HrmC	HrmD	HrmG

densation of a C_5 and a C_3 units to form the eight-carbon skeleton of lincomycin A (Figure 2.2B). However, the exact identities of these two biosynthetic precursors are unknown. The results of feeding experiments suggested that the C_5 unit is likely an intermediate from pentose phosphate pathway, such as D-ribose 5-phosphate (R5P, **2-52**) or D-xylose 5-phosphate (X5P, **2-53**). In addition, the C_3 unit may be derived from S7P (**2-11**) or D-fructose 6-phosphate (F6P, **2-12**) as observed in the transaldolase reaction in the pentose phosphate pathway (Figure 2.2A).

Biosynthesis of the NDP-octose

The fact that several genes homologous to those found in various NDP-deoxyhexose pathways exist in the *lmb* gene cluster prompted us to hypothesize that an NDP-activated octopyranose (**2-57**) is a possible intermediate in the MTL biosynthetic pathway.[115,117] However, this hypothesis lacks literature precedence because NDP-octose has never been reported as an biosynthetic intermediate.

In a typical NDP-sugar biosynthetic pathway, the sugar precursor is first converted to the corresponding sugar 1-phosphate prior to nucleotidyl transfer to yield the NDP-sugar (Figure 2.10). For the biosynthesis of NDP-deoxyhexoses, generation of the hexose 1-phosphate is achieved either by direct C-1 phosphorylation or C-6 phosphorylation followed by a mutase-catalyzed 6→1 migration (**2-64**→**2-65**, Figure 2.10A).[11] A different scenario was observed in the biosynthesis of

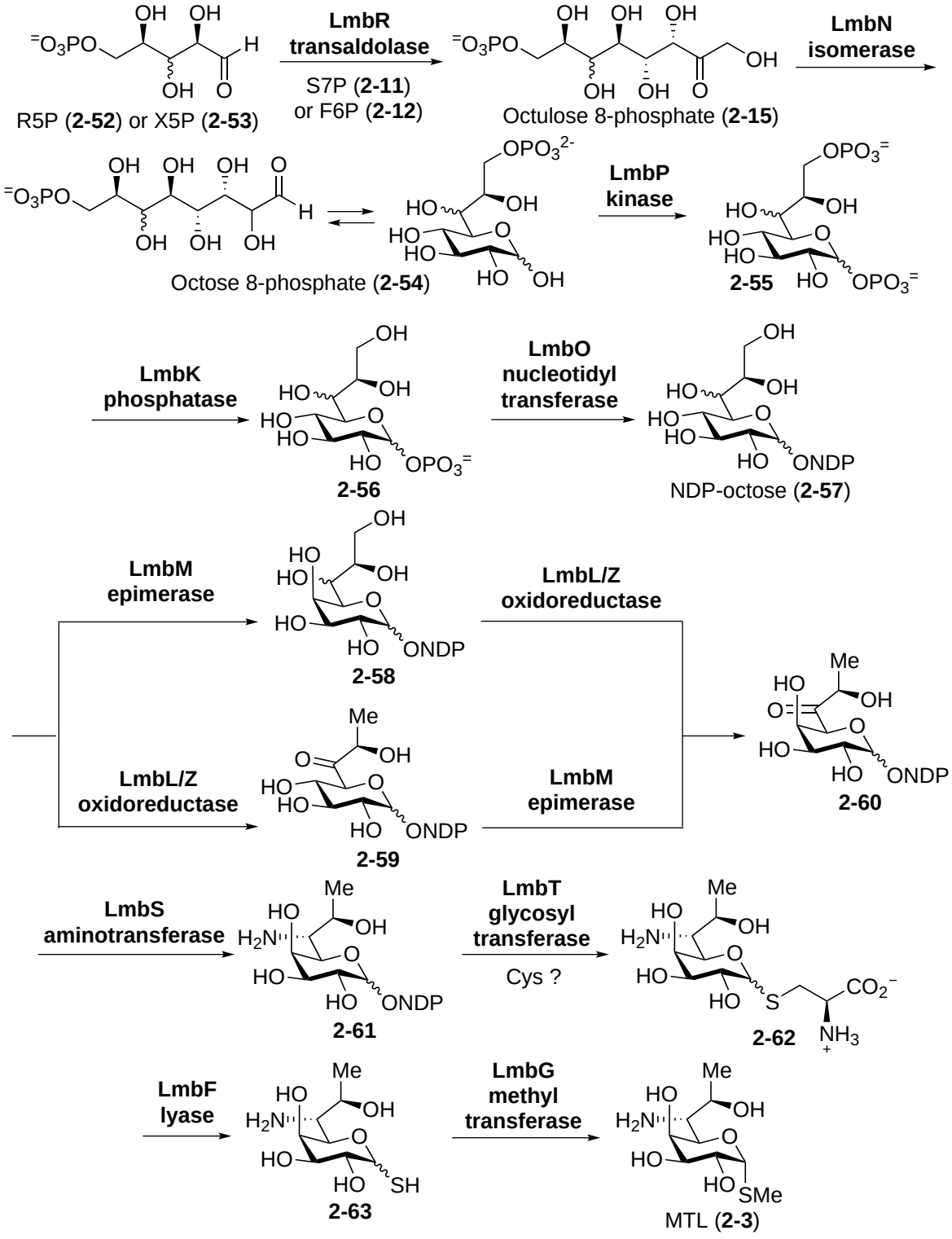


Figure 2.9: Proposed biosynthesis of the MTL moiety

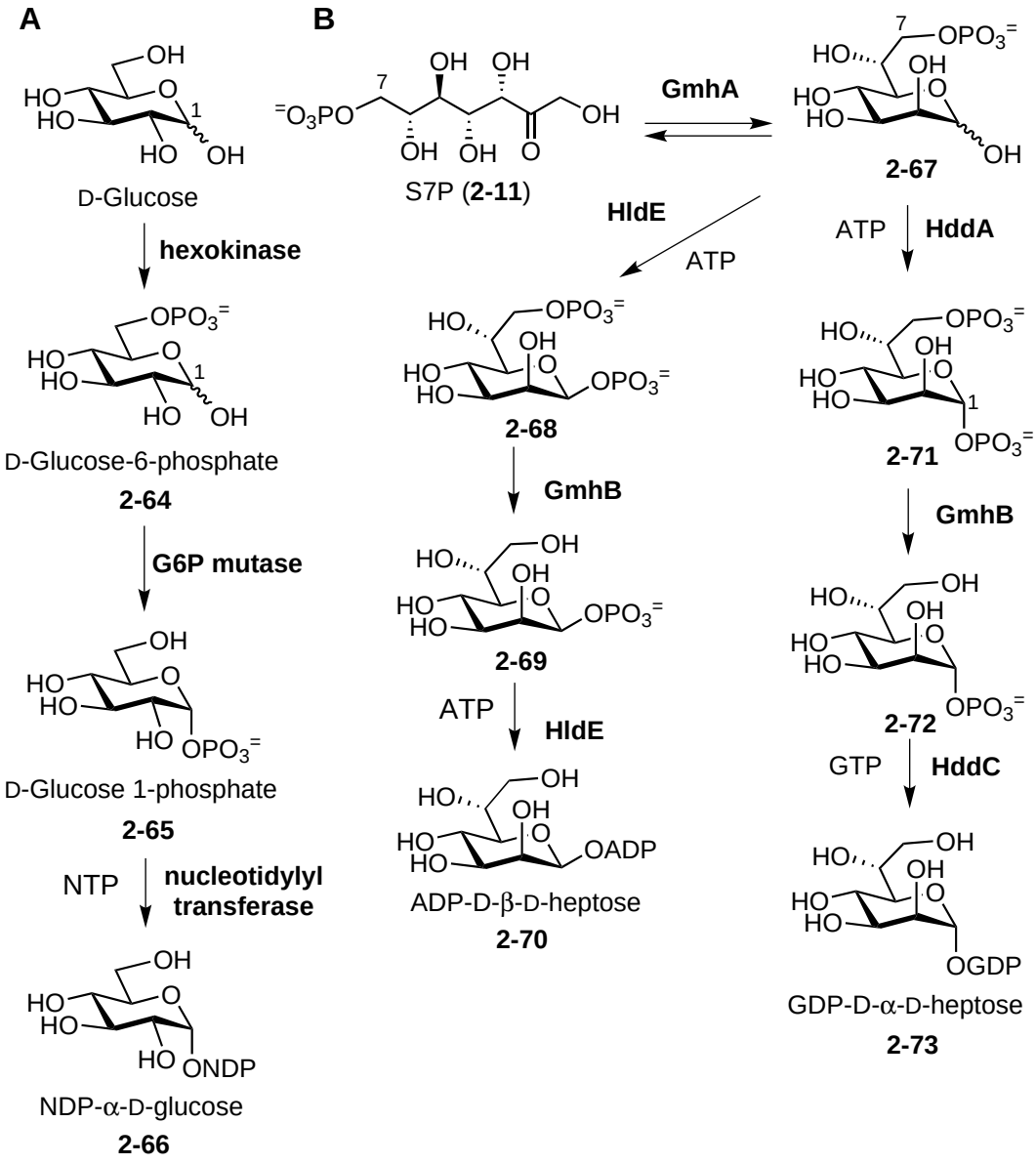


Figure 2.10: Biosynthetic pathways of (A) NDP-hexose and (B) NDP-heptose.

NDP-heptose (Figure 2.10B).[57] Biosynthetic studies of ADP-D-*glycero*-β-D-*manno*-heptose (ADP-D-β-D-heptose, **2-70**) and GDP-D-*glycero*-α-D-*manno*-heptose (GDP-D-α-D-heptose, **2-73**) revealed a kinase/phosphatase cascade to produce heptose 1-phosphate in three steps:

1. Isomerization of S7P to D-*glycero*-D-*manno*-heptose (**2-11**→**2-67**).
2. Anomeric phosphorylation of **2-67** to heptose 1,7-bisphosphate by a kinase (**2-67**→**2-68** and **2-67**→**2-71**).
3. Hydrolysis of the C-7 phosphate group to give heptose 1-phosphate catalyzed by a phosphatase (**2-68**→**2-69** and **2-71**→**2-72**).

Sequence analysis of the *lmb* cluster found no mutase homolog required for installation of the C-1 phosphoryl group via a C_N →C1 migration route. However, comparison of the *lmb* gene cluster with the NDP-heptose biosynthetic gene clusters from *Aneutinibacillus thermocerophilus* DSM 10155/G+[131] and *E. coli* K-12[132] showed several genes, including *lmbN*, *lmbP*, *lmbK* and *lmbO*, whose translated amino acid sequences have similarities to those involved in NDP-heptose biosynthesis (Table 2.7). From this information, it is speculated that MTL biosynthesis may follow a kinase/phosphatase cascade analogous to that observed in the NDP-heptose biosynthetic pathway (Figure 2.10) involving LmbP, LmbK and LmbO to produce the NDP-octose intermediate (**2-57**) (see Figure 2.9).

Table 2.7: Sequence analysis of the genes in lincomycin A and LPS biosyntheses.

	LmbN	LmbP	LmbK	LmbO
GDP-D-α-D-heptose (**2-73**)	GmhA 26 / 43	HddA 23 / 41	GmhB 25 / 42	HddC 27 / 47
ADP-L-β-D-heptose (**2-70**)	GmhA 31 / 47	HldE none	GmhB 32 / 47	HldE none

Enzymatic modifications of NDP-octose

BLAST analysis revealed that *lmbL*, *lmbM* and *lmbS* share sequence homology with those involved in various NDP-sugar pathways (Table 2.4). In addition to the C-1 sulfur transfer, to convert NDP-octose (**2-26**) to MTL, three additional enzymatic modifications can be envisioned: C-4 epimerization, C-6 amination and C-8 dehydration (Figure 2.9). It is thus proposed that LmbM, an epimerase homolog, may catalyze the C-4 epimerization. The C-6 oxidation and C-8 dehydration are possibly catalyzed by a putative NDP-sugar dehydrogenase, LmbL, or a putative oxidoreductase, LmbZ, and the reaction mechanism should be reminiscent to that of TDP-glucose 4,6-dehydratase. However, the order of C-4 epimerization and C-8 dehydration is uncertain. Installation of the amino group at C-6 position catalyzed by an aminotransferase homolog (LmbS) is believed to be the last NDP-octose modification step. The above three modifications are proposed to occur prior to C-1 sulfur transfer, because the translated products of *lmbM*, *lmbL* and *lmbS* share the binding motif recognizing the NDP group.

Proposed pathway of sulfur incorporation at C-1 position

A glycosyltransferase homolog, LmbT, is proposed to incorporate a sulfur donor into the C-1 position of NDP-octose (Figure 2.9). One possible sulfur donor is L-cysteine. To convert the thio-substituent to the free *C*1-thiol intermediate (**2-63**), LmbF, a PLP-dependent aminotransferase homolog, is proposed to cleave the C–S bond of intermediate **2-62**. Finally, a methyltransferase homolog, LmbG, would complete the MTL biosynthesis (Figure 2.9).

An alternative sulfur donor for LmbT is mycothiol (MSH, **2-74**). This hypothesis is based on the finding that LmbE shows moderate sequence similarity to MshB, a *N*-acetyl-1-D-*myo*-inosityl-2-amino-2-deoxy-D-glucopyranoside (GlcNAc-Ins) deacetylase, and Mca, a mycothiol *S*-conjugate amidase (Table 2.4, Figure 2.11).

Both enzymes, MshB and Mca, have been shown to be involved in the MSH biosynthesis or MSH-conjugate catabolism.[133] Therefore, the hypothetic MSH *S*-conjugate (**2-75**) could be hydrolyzed by LmbE to give an intermediate (**2-76**).[133] The subsequent deacetylation might also be catalyzed by LmbE or an unknown cytosolic enzyme to yield the sugar-cysteine conjugate (**2-62**) (Figure 2.11).

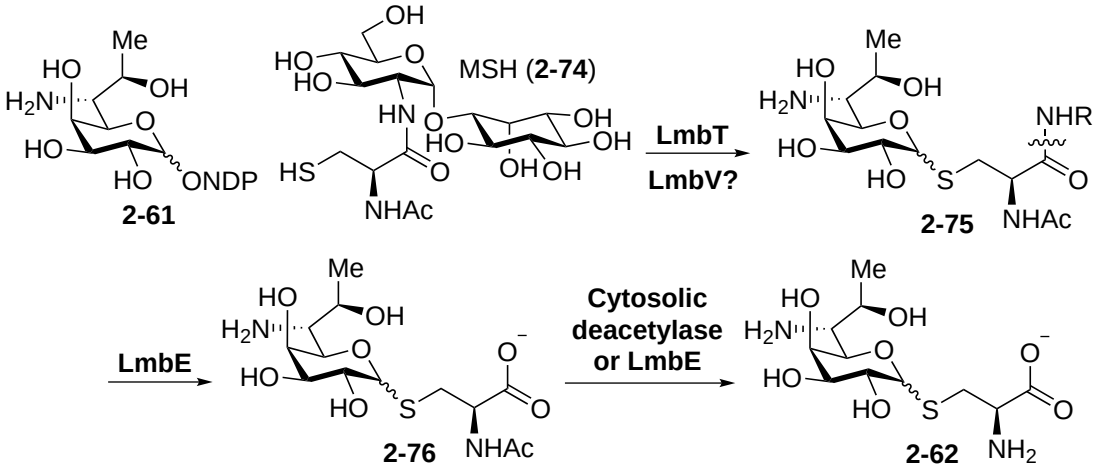


Figure 2.11: Mycothiol as an alternative sulfur donor.

In addition, the protein sequence of LmbV contains a MSH-dependent maleylpyruvate isomerase (MDMPI) domain (Table 2.4). From the results of conserved domains search, LmbV is classified as the member in the DinB superfamily (Din represents the abbreviation of DNA damage-inducible), which contains 8 members so far. Among them, the domains DinB_2 and DUF664 have been found in the known mycothiol and bacillithiol *S*-trasnferases, respectively.[134] If MSH is the sulfur donor, LmbV may serve as an activator enhancing the nucleophilicity of the thiol group and catalyze the sulfur incorporation together with LmbT, which recognizes the NDP-sugar and generates the oxocarbenium intermediate.

Biosynthesis of celesticetin

Based on the sequence analysis shown in Table 2.5 and the discussion of sulfur incorporation above, a modified biosynthetic pathway of SaE moiety (**2-6**) in celesticetin is proposed (Figure 2.12). The sulfur transfer and later transformation steps are speculated to produce intermediate (**2-62**) are believed to be the same as as in lincomycin A biosynthesis (Figure 2.9). However, the fate of the the sugar-cysteine conjugate (**2-62**) may divert due to different reactions carried out by LmbF and CcbF in lincomycin A and celesticetin pathways, respectively. (Figure 2.12). LmbF is annotated as a putative PLP-dependent aminotranferase with 40% identity and 55% similarity to CcbF. While LmbF is proposed to serve as a lyase to cleave the C–S bond (Figure 2.12A), CcbF may execute decarboxylation and aminotransfer to give **2-77** (Figure 2.12B). The resulting aldehyde (**2-77**) can then be further reduced to alcohol (**2-78**) by Ccb5, a putative NADP-dependent alcohol dehydrogenase. Ccb3 and Ccb2 display similarity to salicylate synthase and salycyl-AMP ligase, respectively. Ccb1, annotated as diacylglycerol *O*-acyltransferase, may catalyze the formation of SaE moiety (**2-80**) by ligating **2-78** with salicyl-AMP (**2-79**).

Amide bond formation between the amino acid and sugar moieties

The formation of the amide bond between the carboxylate group of PPL (**2-4**) and the amino group of MTL (**2-3**) is believed to occur in later stages of the biosynthesis of lincomycin A. It has been demonstrated that the cell-free extracts of *S. lincolnensis* could catalyze an ATP-dependent, but ribosome-independent condensation reaction.[135] After fractionization of the crude extract through a size-exclusion column, no single fraction showed the condensation activity, which alternatively could be restored by combining specific fractions. This observation indicated that the nature of the condensation system may be an enzyme complex

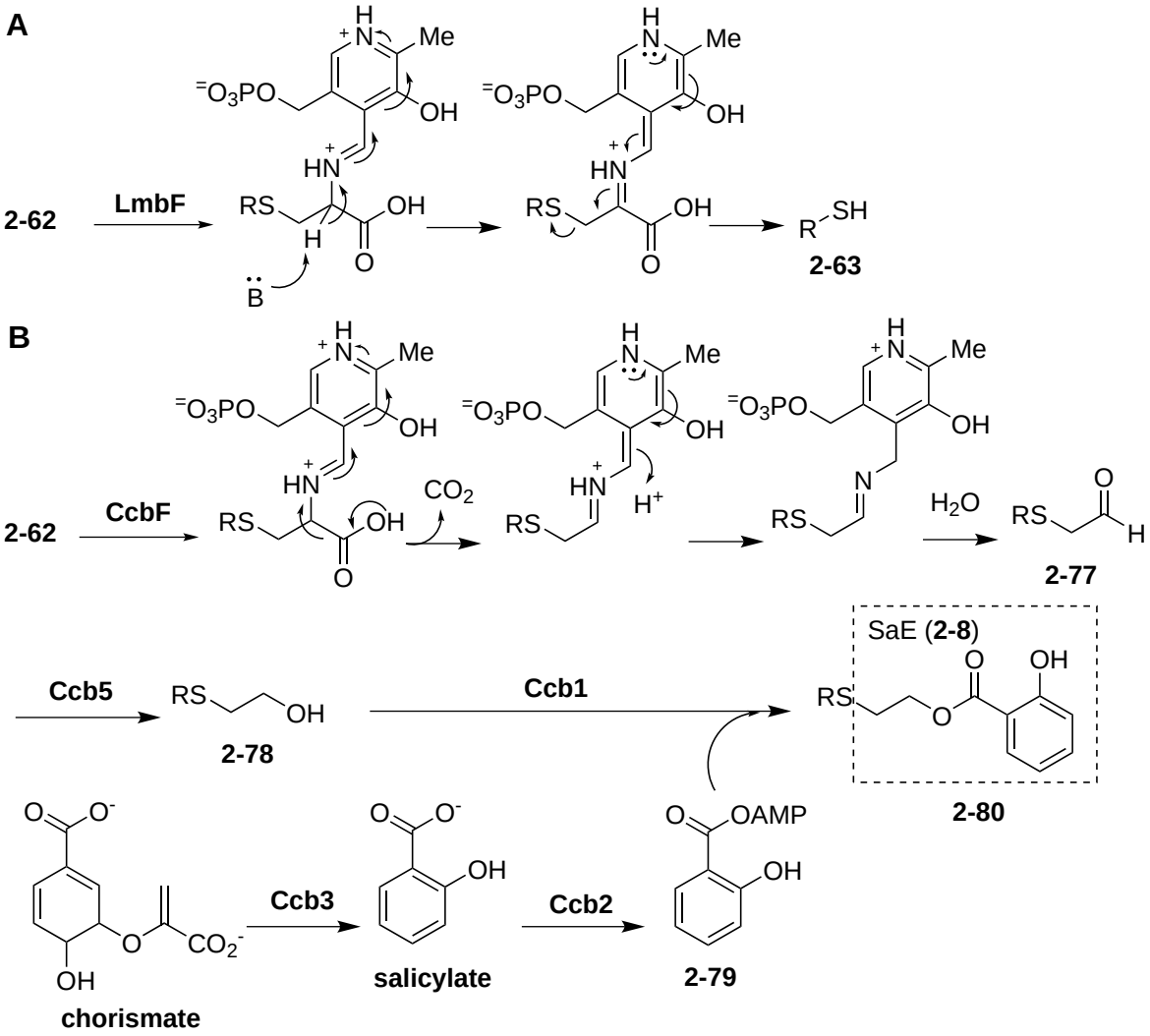


Figure 2.12: Proposed mechanisms catalyzed by (A) LmbF and (B) CcbF.

of several subunits.

Sequence analysis of the lincomycin gene cluster indicates that LmbC/CcbC shows a sequence homology to adenylation domains (A-domains). Adenylation domain is one of the three core domains in a minimal nonribosomal peptide synthetase (NRPS) module.[136] In an assembly line of NRPS, peptide bond formation occurs on an elongating peptidyl chain attached to a carrier protein or thiolation domain (T-domain). The A-domain is responsible for selecting the amino acid monomer and the condensation domain (C-domain) catalyzes the C–N bond formation. The fact that LmbC/CcbC contains an A-domain suggests that they may activate the carboxyl groups of the recognized amino acids (Table 2.4). A recent publication indeed demonstrated that LmbC and CcbC are capable of adenylating propylproline and proline, respectively.[137]

The similarity of the amide bond formation within lincomycin A biosynthesis and NRPS prompts us to search for the missing T- and C-domain in *lmb* biosynthetic gene cluster. The search of homologous sequence of LmbD/CcbD using BLAST in database returned absurd prediction of the enzymatic function. Although these two proteins display high mutual sequence homology (identity / similarity = 56% / 65%, page 56), no significant homology could be found in the known sequences in database. However, the fragments of LmbD actually showed slight homology (>30%) to fusaric acid resistance protein (residue 153-294), NRPS thioesterase (residue 132-306) and UDP-MurNAc-Ala-Glu ligase (residue 230-342). Fusaric acid exhibits a similar chemical structure as that of propylproline, which is the amino acid moiety of lincomycin A. Moreover, the understandings of the enzyme mechanisms of NRPS thioesterase and ligase led to the proposal that LmbD, annotated as hypothetical protein, may serve as the putative C-domain (Figure 2.13A).

Interestingly, the DNA sequence coding for a putative T-domain is found in

the 5′-terminal part of *lmbN* gene in the lincomycin cluster, whereas in celesticetin cluster, the putative T-domain is located in the 3′-terminal part of the *ccbZ* gene, and is right next to the *ccbN* gene (Figure 2.13B). Thus, LmbN may be a bi-functional protein playing different roles in two separate stages of lincomycin A biosynthesis. The *C*-terminal motif is proposed to show isomerization activity in the formation of MTL and the *N*-terminal domain serves as a peptidyl carrier protein (PCP) required for the amino acid condensation.

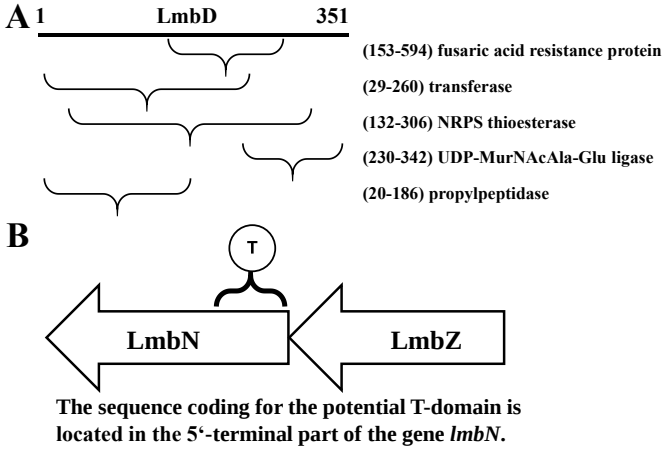


Figure 2.13: (A) Sequence analysis of LmbD. (B) Bi-functional protein LmbN.

2.3.2 Functional Characterization of LmbR

To verify the proposed transalsolase activity of LmbR (Figure 2.2), the recombinant *C*-terminal His_6-tagged LmbR was overexpressed in *E. coli* and purified. The mechanism of transaldolase reaction can be divided into two half reactions (Figure 2.14).[138] The first half reaction involves the formation of a Schiff base intermediate between the C2-keto group of the ketosugar substrate (F6P or S7P) and a lysine residue of the enzyme. Then, the cleavage of C_α–C_β bond via a retro-aldol reaction results in a C_3 unit covalently trapped in the enzyme active site.

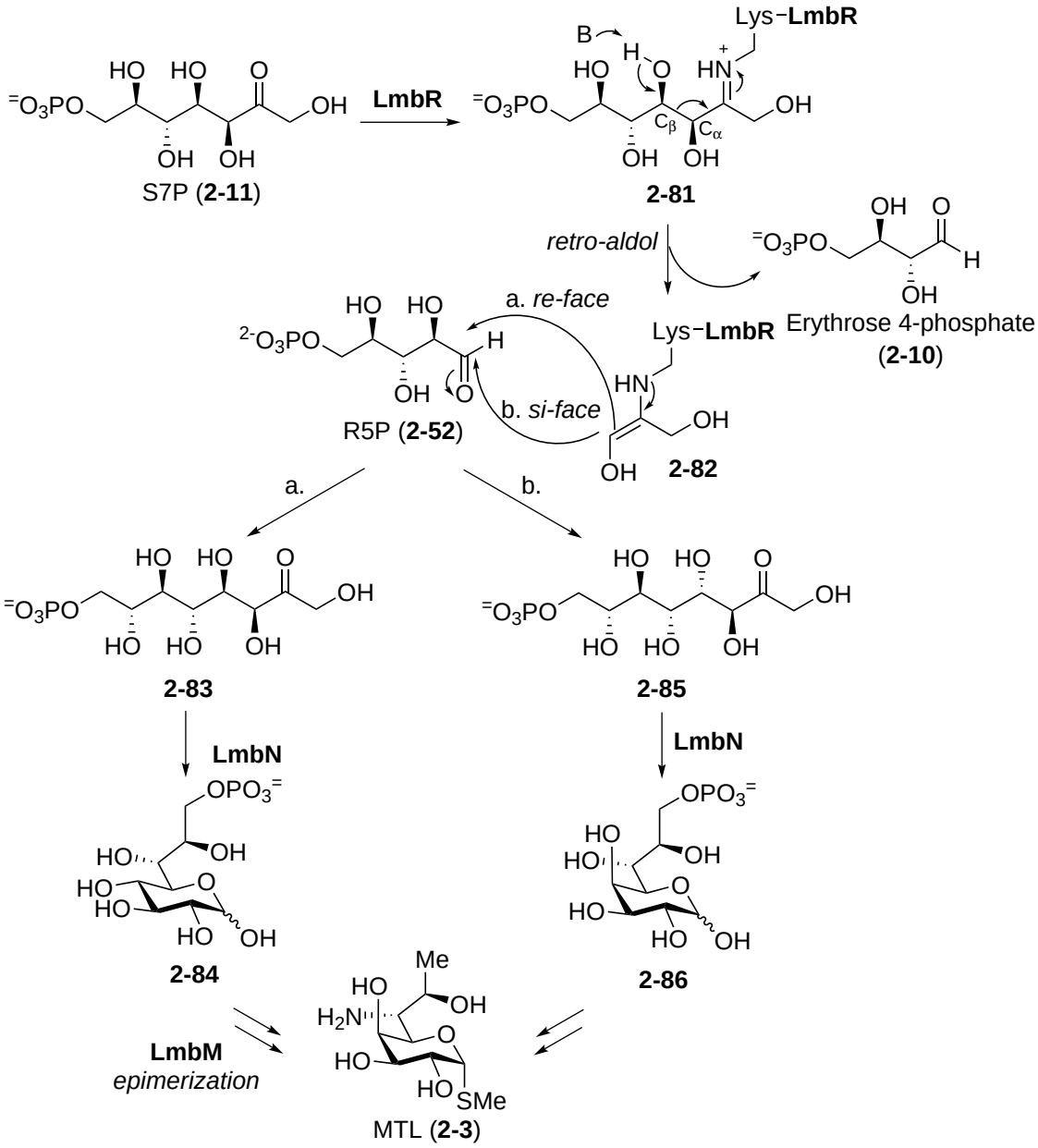


Figure 2.14: Proposed mechanism for LmbR and LmbN-catalyzed reactions.

To examine whether S7P (**2-11**) is a competent C_3 donor for the LmbR-catalyzed reaction, the purified enzyme was incubated with S7P (**2-11**) followed by the reduction with sodium borohydride treatment. The reaction mixture was subjected to MS analysis. Under this condition, MS signals corresponding to the reduced forms of the *C*-His_6-LmbR/S7P conjugate (reduced **2-81**, calcd. 25231 Da; obsd. 25231 Da) and *C*-His_6-LmbR/dihydroxyacetone conjugate (reduced **2-82**, calcd. 25026 Da; obsd. 25030–25045 Da broad peak) were observed in addition to the unmodified enzyme (Figure 2.15). This result indicate that S7P (**2-11**) is a possible substrate of LmbR and the reaction proceeds in a similar manner as other transaldolases involving the formation of an imine adduct followed by C_α–C_β bond cleavage via a retro-aldol reaction (Figure 2.14). The resulting C_3 unit bound in the active site of LmbR was later shown to react with R5P (**2-52**) to yield an octulose 8-phosphate by another graduate student, Eita Sasaki, and the results were published elsewhere.[139]

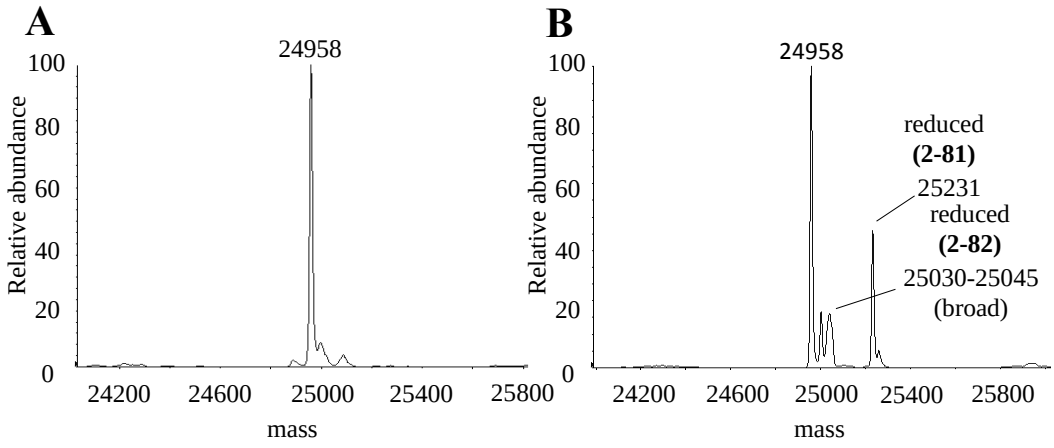


Figure 2.15: Deconvoluted ESI-MS of *C*-His_6-LmbR. (A) *C*-His_6-LmbR treated with $NaBH_4$. (B) *C*-His_6-LmbR in the presence of S7P (**2-11**) treated with $NaBH_4$. Reduced *C*-His_6-LmbR/S7P conjugate (reduced **2-81**, calcd. 25231 Da; obsd. 25231 Da); reduced *C*-His_6-LmbR/dihydroxyacetone conjugate (reduced **2-82**, calcd. 25026 Da; obsd. 25030–25045 Da). The broad peaks were not well-resolved due to the weak signal intensity and the limitation of instrument.

2.3.3 Stereochemical Analysis of LmbN-product

While it has been demonstrated that the recombinant LmbN can isomerize the ketosugar product, resulting from LmbR-catalyzed reaction, to an aldose,[139] the stereochemistry of the C-4 hydroxyl group remains elusive. The sterochemistry uncertainty is due to the stereoselectivity of LmbR-catalyzed transaldol condensation. Based on the general stereochemistry established for most aldolase-catalyzed reaction, the resulting octulose 8-phosphate is expected to inherit the stereochemistry at the C-3 and C-4 positions from the corresponding chiral centers of the C_3 donor (Figure 2.14).[138] This would yield an octose intermediate having C3-(*S*) and C4-(*R*) configuration (**2-83**). However, the C4-(*R*) configuration is in contrast to the C4-(*S*) configuration of the final product, MTL (**2-3**). This is the reason why we proposed a putative C4-epimerase (encoded by the *lmbM* gene) may participate in the later stage of lincomycin A biosynthetic pathway (Figure 2.9). Nevertheless, reactions catalyzed by aldolases giving the inverse sterochemistry are also known.[140] It is thus possible that the LmbR-catalyzed reaction may generate an octose intermediate with C4-(*S*) hydroxyl group (**2-85**).

In order to fully characterize the LmbR and LmbN reactions, the stereochemistry of the C-4 hydroxyl group has to be determined. Although a small amount of enzymatic product could be isolated for MS analysis, it was difficult to secure sufficient amounts for NMR characterization. Thus, we opted to chemically synthesize the peracetylated C4-(*R*)- and C4-(*S*)-octose standars (**2-27** and **2-39**, respectivtly (Figure 2.3 and 2.4). For comparative analysis, the LmbR products were treated with alkaline phosphatase (to remove phosphate groups) and the resulting sugars were peracetylated using acetic anhydride under basic condition. The derivatized enzymatic product mixtures and the synthetic standards were then analyzed by HPLC.

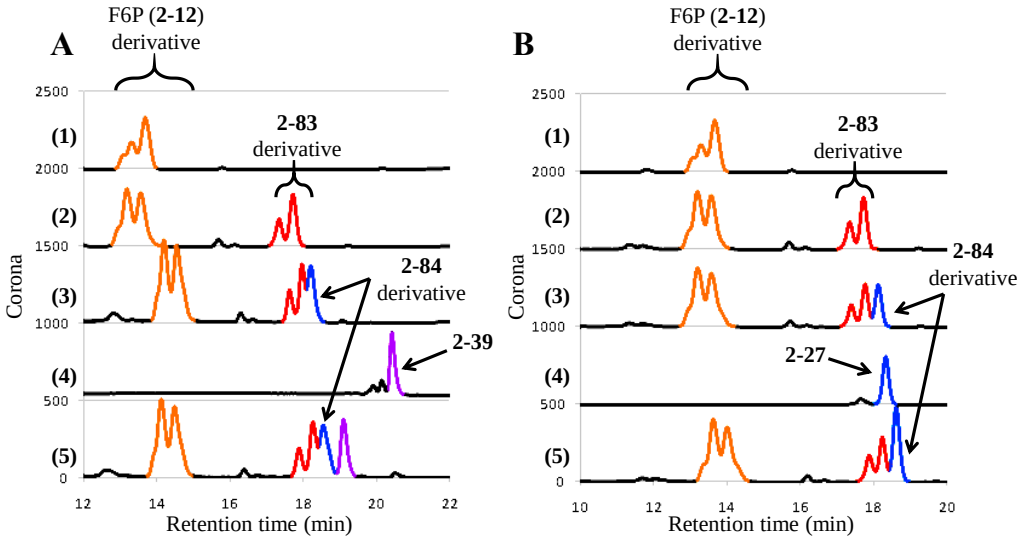


Figure 2.16: (A1) & (B1) Control reaction with F6P (**2-12**) after dephosphorylation and acetylation. (A2) & (B2) LmbR reaction after dephosphorylation and acetylation. (A3) & (B3) LmbR and LmbN reactions after dephosphorylation and acetylation. (A4) Synthetic standard **2-39**. (A5) Coinjection of the sample derived from LmbR and LmbN reaction, and the synthetic standard **2-39**. (B4) Synthetic standard **2-27**. (B5) Coinjection of the sample derived from LmbR and LmbN reaction, and the synthetic standard **2-27**.

As shown in Figure 2.16, two sets of signals (retention times ~13.5 and ~17.5 min) arise from the LmbR reaction (trace A2). The first set of two peaks matches the signals observed for the control reaction, which generate pentaacetylated fructose as the product (trace A1). The splitting peak pattern likely reflects the formation of α and β anomers during chemical derivatization. The second set of two peaks (retention time ~17.5 min, trace A2) can be attributed to the the octulofuranose heptaacetate derived from the LmbR product (**2-83**). When both LmbR and LmbN were used, a new peak with a retention time of ~18 min emerged (trace A3). The MS analysis of this new peak agreed with the peracetylation derivative of the proposed LmbN product (**2-84**). Importantly, the retention time of this peak matches that of the C4-(*R*) standard **2-27** (trace B4). Moreover, this product and **2-27** coeluted when coinjected (trace B5). In contrast, the C4-(*S*) isomer (**2-39**) eluted with a longer retention time (trace A4) than the peracetylated LmbN

product (trace A4). These results unambiguously show that the C-4 position of the LmbR/N reaction products has a (*R*)-configuration (Figure 2.14, route a).

2.4 Conclusion

In this chapter, the biosynthetic pathway of lincomycin A was proposed based on the results from feeding experiments documented in the literature and bioinformatic analysis. Moreover, our biochemical characterization of the two initial enzymatic steps in the MTL biosynthetic pathway led to the identification of an important intermediate, D-*erythro*-D-*gluco*-octose 8-phospahte. Our experiments demonstrated that this intermediate is formed via a transaldol reaction catalyzed by LmbR. Subsequent 1,2-isomerization catalyzed by LmbN could convert the resulting 2-ketose, octulose 8-phosphate, to an aldose, octose 8-phosphate. These results are consistent with our proposed biosynthetic pathway of lincomycin A (Figure 2.9). The HPLC analyses using the chemically synthesized peracetylated octoses also supported the predicted functions of enzymes. More importantly, this piece of information provides, for the first time, *in vitro* evidence for the biosynthetic origin of the eight-carbon backbone of MTL, and lays a solid foundation for later biosynthetic studies of lincomycin A.

Chapter 3

Biosynthetic Studies of Lincomycin A (II): Formation of the GDP-Octose Intermediate*

3.1 Introduction

Having confirmed the biosynthetic origin of the thiooctose moiety in lincomycin A biosynthesis firmly lays a cornerstone supporting our proposed pathway. In addition to the identification of octose 8-phosphate intermediate, the characterization of LmbR and LmbN integrates the results of early feeding experiments and updated bioinformative analysis. This piece of evidence, in turns, allows us to pursue further investigation for lincomycin A biosynthesis.

Lincomycin A contains a highly modified sugar moiety, which is likely fabricated through intensive enzymatic transformations of an unmodified octose precursor. Moreover, the sulfur incorporation is predicted to proceed in a similar manner as the assembly of activated sugars with aglycon substrates catalyzed by glycosyltransferase. In order to covert the unmodified octose to an intermediate which could be used by the biosynthetic enzymes and glycosyltransferases, the monosaccharaide must be activated as a nucleoside diphosphate (NDP) derivative. In the biosynthesis of secondary metabolites, activation usually involves the coupling of a sugar 1-phosphate to an nucleoside monophosphate (NMP) moiety

*Partial content of this chapter was published as "*In vitro* Characterization of LmbK and LmbO: Identification of GDP-D-*erythro*-D-*gluco*-Octose as a Key Intermediate in Lincomycin A Biosynthesis" Lin,C.-I; Sasaki, E.; Zhong, A.; Liu, H.-w., *J. Am. Chem. Soc.*, **2014**, *136*, 906–909. The experiments were designed and performed by C.-I L. One of the enzymes was prepared by E. S. One of the substrates was synthesized by C.-I L. with the help from Z. A. This research project was supervised by H.-w. Liu.

from the corresponding nucleoside triphosphate (NTP) by a nucleotidylyltransferse.[141]

Based on the above-mentioned information and the sequence analysis of the *lmb* gene cluster discussed in Chapter 2, we speculated that a NDP-octose intermediate manufactured via a kinase/phosphatase cascade may participate in the biosynthesis of lincomycin A (Figure 3.1). After LmbN catalyzes the isomerization to produce the corresponding octose 8-phosphate (**3-1**),[139] the putative kinase, LmbP, may phosphorylate the C-1 hydroxyl group of octose 8-phosphate (**3-1**) to afford octose 1,8-bisphosphate (**3-2**). The C-8 phosphate group of this bisphosphate intermediate might be hydrolyzed by LmbK (annotated as phosphatase) to give octose 1-phosphate (**3-3**), which is then converted to the nucleotide-activated octose (**3-5**) by LmbO (annotated as nucleotidylyltransferase) (route a). It is also conceivable that the reaction proceeds first with nucleotidylyltransfer by LmbO (**3-2** $\rightarrow$ **3-4**) followed by LmbK-catalyzed C-8 dephosphorylation (**3-4** $\rightarrow$ **3-5**) (route b). Because LmbP and LmbO exhibit sequence similarities only to their counterparts in the GDP-D-α-D-heptose biosynthetic pathways (Table 2.7 in page 63), the C-1 activated octose intermediate in lincomycin biosynthesis is likely a GDP-octose. However, it is well known that prediction of substrate specificity for enzymes based solely on sequence alignment can be erroneous,[10] and thus the chemical nature of sugar phosphodinucleotide intermediate in the lincomycin A biosynthetic pathway must be experimentally verified.

3.2 Experimental Procedures

3.2.1 General

Materials

All chemicals and reagents were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA) and were used without

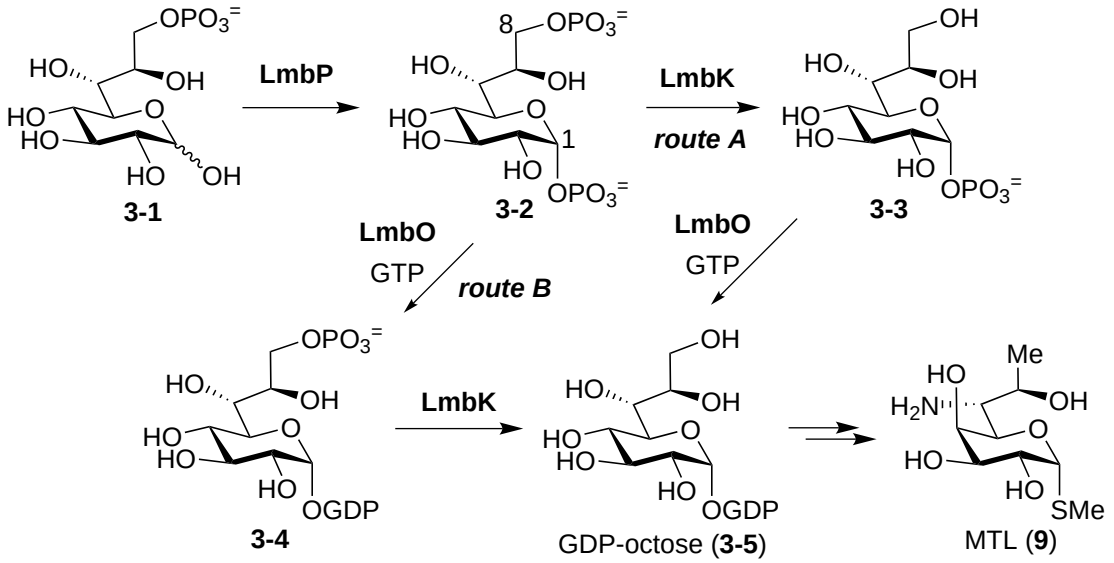


Figure 3.1: Proposed pathway for NDP-octose formation in MTL biosynthesis.

further purification unless otherwise specified. Oligonucleotide primers were prepared by Integrated DNA Technologies (Coralville, IA). Kits for DNA gel extraction and spin minipreps were purchased from Qiagen (Valencia, CA). PureLink® Genomic DNA Mini Kit was bought from Invitrogen (Carlsbad, CA). KOD NDA polymerase was acquired from Novagen (Madison, WI). Enzymes and molecular weight standards used for the cloning experiments were obtained from New England Biolabs (Ipswich, MA). Reagents for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) were purchased from Bio-Rad (Hercules, CA), with the exception of the protein molecular weight markers, which were obtained from Invitrogen. Growth medium components were acquired from Becton Dickinson (Sparks, MD). Sterile syringe filters are products of Fisher Scientific. Amicon YM-10 ultrafiltration membranes were bought from Millipore (Billerica, MA). The CarboPac PA1 high-performance liquid chromatography (HPLC) columns were obtained from Dionex (Sunnyvale, CA). Analytical C_{18} HPLC columns are products of Varian (Palo Alto, CA). Simi-preparative C18 HPLC columns were purchased from Fisher Scientific (Pittsburgh, PA).

Bacterial Strains and Plasmids

Streptomyces lincolnensis NRRL ISP-5355 (identical to ATCC 25466) and *Streptomyces caelestis* NRRL 2418 were obtained from the Agricultural Research Service (ARS) Culture Collection of the National Center for Agricultural Utilization Research (Peoria, IL). *Escherichia coli* DH5α, acquired from Bethesda Research Laboratories (Gaithersburg, MD), was used for routine cloning experiments. The protein overexpression host *E. coli* BL21 star (DE3) was obtained from Invitrogen. Vectors pET28b(+) and pET24b(+) for protein overexpression was purchased from Novagen (Madison, WI).

Instrumentation

Standard genetic manipulations of *E. coli* were performed as described by Sambrook and Russell.[142] DNA sequencing was conducted at the core facility of the Institute of Cellular and Molecular Biology, The University of Texas at Austin. Vector NTI Advance 10.1.1 from Invitrogen was used for sequence alignments. DNA concentrations were measured using a NanoDrop ND-1000 UVvis instrument from Thermo Fisher Scientific. HPLC was performed on a Beckman Coulter System Gold HPLC equipped with a UV detector or a Corona CAD (charged aerosol detector, ESA Biosciences, Chelmsford, MA). Analytical thin layer chromatography (TLC) was carried out on precoated TLC glass plates (silica gel, grade 60, F254, 0.25 mm layer thickness) purchased from EMD chemicals (Madison, WI). Flash column chromatography was performed on silica gel (40–63 μm, 60 Å) obtained from SiliCycle Inc (Quebec City, Canada). NMR spectra were acquired using a Varian Unity 400 MHz, 500 MHz or 600 MHz spectrometer, and chemical shifts (in ppm) are reported relative to that of the solvent peak (δ_H = 7.24 and δ_C = 77.0 for deuterated chloroform, δ_H = 4.67 for deuterium oxide). Mass spectra were recorded at the Mass Spectrometry core facility in the Department of Chem-

istry and the Proteomics Facility in College of Pharmacy at the University of Texas, Austin. Optical rotations were measured using an ATAGO AP-300 automatic polarimeter at a path length of 1 dm.

3.2.2 Genomic DNA Extraction

S. lincolnensis

Bacterial cells of *S. lincolnensis* NRRL ISP-5355 (identical to the ATCC 25466 strain) in a dormant state were inoculated into 7 mL of the International Streptomyces Project (ISP) medium 2 and grown in a rotary incubator at 30 °C and 250 rpm for 16 h. The resultant seed culture (1 mL) was transferred to 50 mL of ISP-2 medium and grown under the same conditions for 2 days. The mycelia were harvested by centrifugation at 5000×g for 10 min. The cells were resuspended in 10 mL of TES buffer (25 mM Tris, 25 mM EDTA, and 10.3% sucrose, pH 8) and treated with lysozyme (15 mg) at 37 °C and 250 rpm for 1 h. To the resulting mixture was added SDS (final concentration 0.3%) and proteinase K (1.5 mg), and the combined solution was incubated under the same conditions for an additional 1 h. The released genomic DNA and RNA mixture was isolated through phenol-chloroform extraction followed by sodium acetate-isopropyl alcohol precipitation. The resulting solution was treated with ribonuclease (RNase), and the genomic DNA was isolated by another round of phenol-chloroform extraction and sodium acetate-isopropyl alcohol precipitation. The final concentration of the genomic DNA obtained was 160 ng/μL in 200 μL of 10 mM Tris · HCl buffer (pH 8.5).

S. caelestis

The freeze-dried sample of *S. caelestis* NRRL 2418 bacterial cells were inoculated into 7 mL of the tryptic soy broth (TSB) medium and grown in a rotary incubator at 30 °C and 250 rpm for 24 h. The resultant seed culture (1 mL) was

transferred to 50 mL of TSB medium and grown under the same conditions for 1 day. The overnight culture (5 mL) was harvested by centrifugation at 4,500×g for 10 min. The cells were resuspended in 360 µL of lysozyme digestion buffer (25 mM Tris · HCl pH 8, 2.5 mM EDTA, 1% Triton X-100) and treated with lysozyme (final concentration 20 mg/mL) at 37 °C for 30 min. To the resulting mixture was added 40 µL of proteinase K solution and 400 µL of PureLink™ Genomic Lysis/Binding buffer supplied with the PureLink® Genomic DNA kit. The combined solution was incubated at 55 °C for 30 min. To the lysate was added 400 µL 96-100% ethanol to yield a homogenous solution. The lysate (∼ 1.28 mL) was transferred to the PureLink™ Spin column and centrifuged at 10,000×g for 1 min. The flow-through was discarded. The spin column was washed with 500 µL PureLink™ Wash buffer 1 (1×) and 500 µL PureLink™ Wash buffer 2 (1×) by centrifugation at 10,000×g for 1 min each. The spin column was then placed in a sterile 1.5-mL microcentrifuge tube and the DNA was eluted with 100 µL of 10 mM Tris · HCl buffer (pH 8.5). The final concentration of the genomic DNA obtained was 60.4 ng/µL.

3.2.3 Cloning of *lmbP*, *lmbK*, *lmbO*, *ccbP* and *ccbO*

Plasmids containing *lmbP*, *lmbK*, *lmbO*, *ccbP* and *ccbO* were constructed by amplifying the corresponding genes from genomic DNA isolated from *S. lincolnensis* and *S. caelestic* with designed primer pairs (Table 3.1). The PCR products were then digested with NdeI and HindIII and ligated into pET24b(+) and pET28b(+), which had been digested with the same restriction enzymes. The resulting plasmids were sequenced using T7 or T7 terminal universal primer and used to transform the *E. coli* BL21 star (DE3) for protein overexpression.

Table 3.1: Primers used for constructing plasmids containing *lmbP*, *lmbK*, *lmbO*, *ccbP* and *ccbO* genes.

Plasmid	Primer pair
pET24b(+)-*lmbP*-*C*His$_6$	CIL001 / CIL002
pET28b(+)-*lmbP*-*N*His$_6$	CIL001 / CIL003
pET24b(+)-*lmbP*	CIL001 / CIL003
pET28b(+)-*lmbK*-*N*His$_6$	CIL004 / CIL005
pET24b(+)-*lmbO*-*C*His$_6$	CIL006 / CIL007
pET28b(+)-*lmbO*-*N*His$_6$	CIL006 / CIL008
pET24b(+)-*lmbO*	CIL006 / CIL008
pET24b(+)-*ccbP*-*C*His$_6$	CIL009 / CIL010
pET28b(+)-*ccbP*-*N*His$_6$	CIL009 / CIL011
pET24b(+)-*ccbP*	CIL009 / CIL011
pET24b(+)-*ccbO*-*C*His$_6$	CIL012 / CIL013
pET28b(+)-*ccbO*-*N*His$_6$	CIL012 / CIL014
pET24b(+)-*ccbO*	CIL012 / CIL014

Primer number	Primer name	Sequence[a]
CIL001	F-*lmbP*-24/28-NdeI	5′-ATCACTA**CATATG**ATCGACGTCACGGCGCC-3′
CIL002	R-*lmbP*-24-HindIII	5′-TCAT**AAGCTT**TGCCCGCCCCTCTCCG-3′
CIL003	R-*lmbP*-28-HindIII	5′-TCAT**AAGCTT***TCA*TGCCCGCCCCTCTCC-3′
CIL004	F-*lmbK*-24/28-NdeI	5′-ATGTGAA**CATATG**GGACGCGAGGGACAGTCG-3′
CIL005	R-*lmbK*-28-HindIII	5′-AGAT**AAGCTT***TCA*GCGGCCGCCGTGCCCG-3′
CIL006	F-*lmbO*-24/28-NdeI	5′-ATCACTA**CATATG**GTCGCGTCGACCGAACC-3′
CIL007	R-*lmbO*-24-HindIII	5′-ATCT**AAGCTT**TTTCTCACCTGTCCGCAGATAGC-3′
CIL008	R-*lmbO*-28-HindIII	5′-TCAT**AAGCTT**CGA*TCA*TTTCTCACCTGTCCGC-3′
CIL009	F-*ccbP*-24/28-NdeI	5′-TCTATAA**CATATG**ATCACGACCACCGCTCC-3′
CIL010	R-*ccbP*-24-HindIII	5′-ATAT**AAGCTT**-CGCTGTGCGCACCCCG-3′
CIL011	R-*ccbP*-28-HindIII	5′-TTAC**AAGCTT***CA*CGCTGTGCGCACC-3′
CIL012	F-*ccbO*-24/28-NdeI	5′-TCTATAA**CATATG**GTTGCGCTGCCCGG-3′
CIL013	R-*ccbO*-24-HindIII	5′-ATAT**AAGCTT**-TATCGTCTCCGCGAATCG-3′
CIL014	R-*ccbO*-28-HindIII	5′-TTAC**AAGCTT***CA*TATCGTCTCCGCGAATCG-3′

[a]The engineered restriction sites are shown in bold; the start codon is shown in bold and underlined; the stop codon is shown in italic.

3.2.4 Protein Overexpression in *E. coli*

A typical procedure for the overexpression of the *N*-His$_6$- or *C*-His$_6$-tagged proteins was performed. An overnight culture of *E. coli* BL21 star (DE3) transformant, grown in the LB medium (10 mL) containing 30 μg/mL of kanamycin at 37 °C, was used to inoculate 1 L of the same growth medium. The culture was in-

cubated at 37 °C with shaking (200 rpm) until the OD_{600} reached ∼0.5. Protein expression was then induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM, and the cells were allowed to grow at 37 °C and 125 rpm for an additional 24 h. The cells were harvested by centrifugation at 4000×*g* for 15 min and stored at −80 °C until lysis.

3.2.5 Purification of *N*-His_6-LmbK

All purification steps were carried out at 4 °C using Ni-NTA resin according to the manufacturer's protocol with minor modifications. Specifically, the thawed cells (∼5 g) were resuspended in the lysis buffer (20 mL) containing 50 mM Tris · HCl (pH 8), 10% (v/v) glycerol and 10 mM imidazole. After incubation with lysozyme (20 mg) for 30 min, the cells were disrupted by sonication using 10 × 10-s pulses with a 20-s cooling pause between each pulse. The resulting lysate was centrifuged at 20,000×*g* for 20 min, and the supernatant was subjected to Ni-NTA chromatography. Bound protein was eluted using 250 mM imidazole buffer containing 10% glycerol. The collected protein solution was dialyzed against 3 × 1-L of 50 mM Tris · HCl buffer (pH 8) containing 300 mM NaCl and 15% glycerol. The protein solution was then flash-frozen in liquid nitrogen and stored at −80 °C until use. Protein concentration was determined by the Bradford assay using bovine serum albumin as the standard.[143] The yield of *N*-His_6-LmbK was approximately 40 mg from 1 L culture. The molecular mass and purity of *N*-His_6-LmbK were determined by SDS-PAGE analysis.

3.2.6 Refolding of *N*-His_6-LmbO from Inclusion Bodies

After the cells were harvested, the cells (∼10 g) were resuspended in 40 mL of the lysis buffer containing 50 mM Tris · HCl (pH 8), 10% glycerol and 10 mM imidazole. After incubation with lysozyme (40 mg) for 1 h, the cells were disrupted

by sonication using 10 × 10-s pulses with a 20-s cooling pause between each pulse. The cell lysate was centrifuged at 20,000×g for 20 min, and the supernatant was discarded. The resulting pelleted inclusion bodies were washed three times with 10 mL of lysis buffer and dissolved in 15 mL of 10 mM Tris · HCl buffer (pH 8) containing 6 M guanidine and 100 mM NaH_2PO_4. The solution was incubated with shaking at 37 °C for 1 h and centrifuged at 20,000×g for 20 min to remove insoluble materials. The solublized *N*-His_6-LmbO inclusion bodies (∼60 mg) were then added dropwise to 40 mL of the refolding buffer composed of 100 mM Tris · HCl buffer (pH 8) containing 20% glycerol, 1 M urea, 50 mM NaCl, 400 mM arginine, and 10 mM dithiothreitol (DTT) with vigorous stirring. This mixture was kept at 4 °C with gentle stirring for 20 hr. The collected protein solution was dialyzed against 1 L of 50 mM Tris · HCl buffer (pH 8) containing 20% glycerol, 300 mM NaCl and 1 mM DTT. Afterward, the protein was centrifuged to remove precipitated material and the supernatant was flash-frozen in liquid nitrogen and stored at −80 °C until use. Protein concentration was determined by the Bradford assay using bovine serum albumin as the standard.[143] The molecular mass and purity of the protein, *N*-His_6-LmbO, were determined by SDS-PAGE analysis.

3.2.7 Synthesis of D-*erythro*-D-*gluco*-octose α-1,8-bisphosphate (3-2)

Methyl 2,3,4-tri-*O*-benzyl-7,8-dideoxy-D-*glycero*-α-D-*gluco*-oct-7-enopyranoside (3-9).

The syntheses of compounds **3-7**, **3-8** and **3-9** as illustrated in Figure 3.2 follow the previously described procedures in Section 2.2.4.

Methyl 2,3,4,6-tetra-*O*-benzyl-7,8-dideoxy-D-*glycero*-α-D-*gluco*-oct-7-enopyrano -side (3-11).

Compound **3-9** (0.89 g, 1.81 mmole) was dissolved in anhydrous DMF (10 mL) and cooled to 0 °C. The solution was then treated with sodium hydride (60%, 0.25

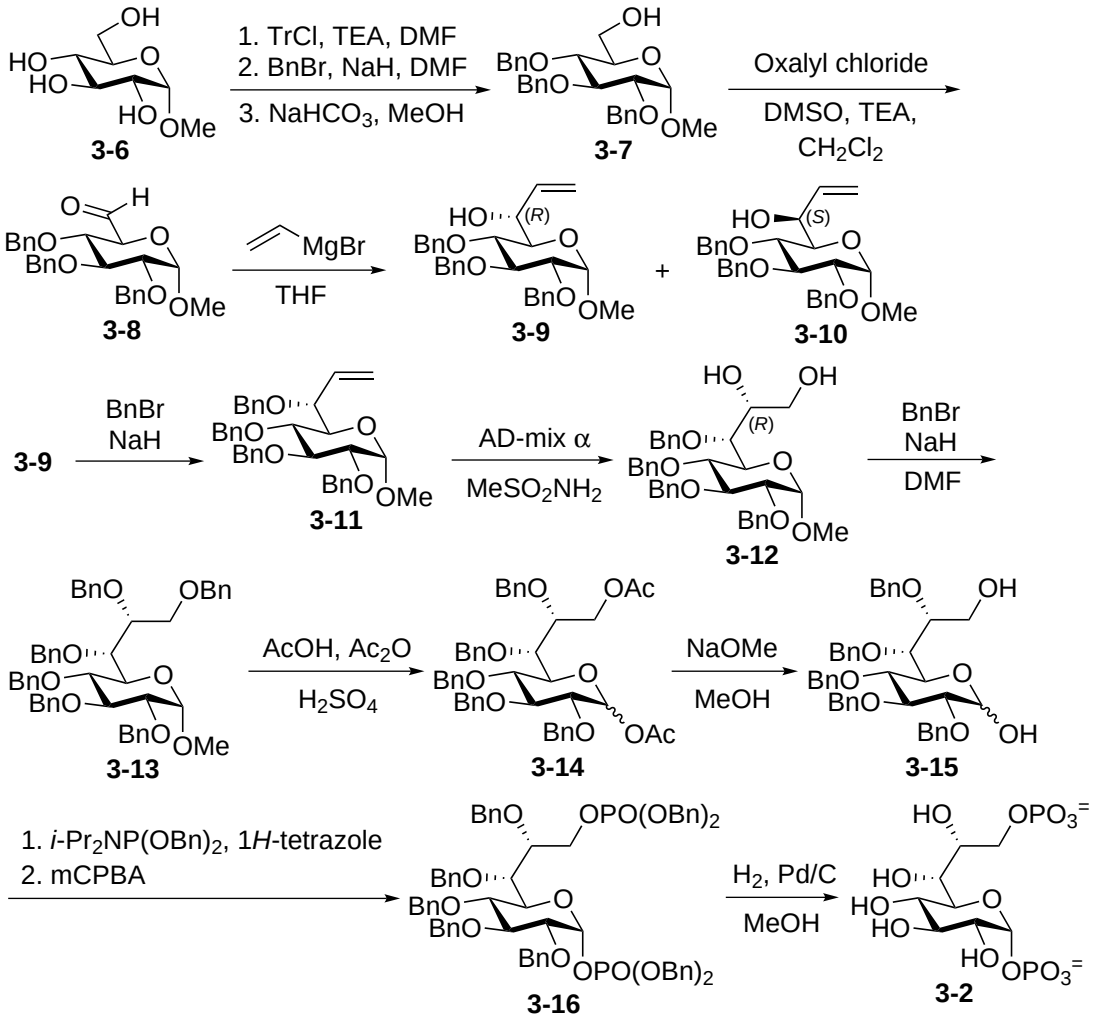


Figure 3.2: Synthesis of octose 1,8-bisphosphate (**3-2**)

g, 6.25 mmole) portion-wise and stirred at 0 °C for 15 min. Benzyl bromide (0.35 mL, 2.94 mmole) was added slowly and the reaction mixture was brought to room temperature. After 16 h, the reaction mixture was quenched by adding methanol slowly and concentrated in vacuo. The resulting residue was mixed with water and extracted with ethyl acetate (3×). The combined organic layers were dried over Na_2SO_4. After the volatiles were removed by rotary evaporation under reduced pressure, the crude product was purified using flash chromatography (silica gel; EtOAC/hexanes (1:9 → 1:4)) to yield **3-11** ($C_{37}H_{40}O_6$, 0.79 g, 75%) as syrup. TLC (EtOAc/hexanes (1:2)) R_f = 0.61. ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) 7.29–6.99 (20 H, m, Bn×4), 5.83 (1 H, ddd, $J_{6,7}$ = 8.8 Hz, $J_{7,8cis}$ =10.4 Hz, $J_{7,8trans}$ = 17.2 Hz, H7), 5.25 (1 H, dd, $J_{8cis,8trans}$ = 1.2 Hz, $J_{7,8cis}$ = 10.4 Hz, H_{8cis}), 5.00 (1 H, dd, $J_{8cis,8trans}$ = 1.2 Hz, $J_{7,8trans}$ = 17.2 Hz, H_{8trans}), 4.92 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.76 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.70 (2 H, J = 11.2 Hz, CH_2 of Bn), 4.59 (1 H, J = 12.0 Hz, CH_2 of Bn), 4.59 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.39 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.34 (1 H, J = 12.0 Hz, CH_2 of Bn), 4.34 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.02 (1 H, dd, $J_{5,6}$ = 1.6 Hz, $J_{6,7}$ = 8.8 Hz, H6), 3.96 (1 H, dd, $J_{2,3}$ = 9.6 Hz, $J_{3,4}$ = 9.2 Hz, H3), 3.91 (1 H, dd, $J_{5,6}$ = 1.6 Hz, $J_{4,5}$ = 10.4 Hz, H5), 3.40 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{2,3}$ = 9.6 Hz, H2), 3.36 (3 H, s, OMe), 3.28 (1 H, dd, $J_{3,4}$ = 9.2 Hz, $J_{4,5}$ = 10.4 Hz, H4). ^{13}C NMR ($CDCl_3$, 100 MHz) δ (ppm) 138.58, 138.23, 138.12, 138.03, 133.83, 129.39, 128.38, 128.34, 128.31, 128.24, 127.99, 127.90, 127.82, 127.52, 127.33, 120.60, 97.55, 82.25, 80.03, 78.79, 78.07, 75.68, 74.41, 73.19, 72.34, 70.01, 54.93. ESI-HRMS Calcd. for $C_{37}H_{40}O_6Na^+$ $[M + Na]^+$ 603.2717, found 603.2725.

Methyl 2,3,4,6-tetra-*O*-benzyl-α-D-*erythro*-D-*gluco*-octopyranoside (3-12).

Compound **3-11** (0.79 g, 1.36 mmole) was suspended in 20 mL of co-solvent ^{t}BuOH : H_2O (1:1). The mixture was cooled to 4 °C in the cold room and added AD-mix α (2.0 g) and methanesulfonamide (0.13 g, 1.37 mmole). The reaction mixture was stirred at 4 °C until the starting material was all consumed as judged by TLC analy-

sis. The reaction was quenched by adding sodium metabissulfite and then concentrated *in vacuo*. The resulting residue was dissolved in ethyl acetate. The organic layer was washed with water (1×) and brine (1×), and then dried over Na_2SO_4. After the volatiles were removed by rotary evaporation, the crude product was purified using flash chromatography (silica gel; EtOAC/hexanes (1:2→2:1)) to yield **3-12** ($C_{37}H_{42}O_8$, 0.67 g, 80%) as syrup. TLC (EtOAc/hexanes (1:1)) R_f = 0.38. ^{1}H NMR ($CDCl_3$, 600 MHz) δ (ppm) 7.37–7.24 (20 H, m, Bn×4), 5.02 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.97 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.82 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.78 (1 H, J = 12.0 Hz, CH_2 of Bn), 4.73 (1 H, J = 11.4 Hz, CH_2 of Bn), 4.67 (1 H, J = 12.0 Hz, CH_2 of Bn), 4.66 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.64 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.48(1 H, J = 11.4 Hz, CH_2 of Bn), 4.06 (1 H, dd, $J_{5,6}$ = 1.2 Hz, $J_{4,5}$ = 10.8 Hz, H5), 4.03 (1 H, dd, $J_{2,3}$ = 9.6 Hz, $J_{3,4}$ = 9.0 Hz, H3), 3.82 (1 H, dd, $J_{3,4}$ = 9.0 Hz, $J_{4,5}$ = 10.8 Hz, H4), 3.79 (1 H, ddd, $J_{6,7}$ = 8.4 Hz, $J_{7,8a}$ = 3.6 Hz, $J_{7,8b}$ = 5.4 Hz, H7), 3.69 (1 H, dd, $J_{5,6}$ = 1.2 Hz, $J_{6,7}$ = 8.4 Hz, H6), 3.67 (1 H, dd, $J_{7,8a}$ = 3.6 Hz, $J_{8a,8b}$ = 11.4 Hz, $H8_a$), 3.55 (1 H, dd, $J_{7,8b}$ = 5.4 Hz, $J_{8a,8b}$ = 11.4 Hz, $H8_b$), 3.53 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{2,3}$ = 9.6 Hz, H2), 3.38 (3 H, s, OMe). ^{13}C NMR ($CDCl_3$, 150 MHz) δ (ppm) 138.58, 138.13, 137.87, 128.52, 128.50, 128.49, 128.47, 128.44, 128.42, 128.39, 128.35, 128.05, 128.01, 127.96, 127.95, 127.93, 127.89, 127.78, 127.66, 97.86, 82.65, 80.22, 79.11, 77.83, 75.79, 75.06, 71.29, 70.34, 55.25. ESI-HRMS Calcd. for $C_{37}H_{42}O_8Na^+$ $[M + Na]^+$ 637.2772, found 637.2777.

Methyl 2,3,4,6,7,8-hexa-*O*-benzyl-α-D-*erythro*-D-*gluco*-octopyranoside (3-13). Compound **3-12** (0.45 g, 0.73 mmole) was dissolved in anhydrous DMF (10 mL) and cooled to 0 °C. The solution was then treated with sodium hydride (60%, 0.10 g, 2.50 mmole) portion-wise and stirred at 0 °C for 15 min. Benzyl bromide (0.2 mL, 1.68 mmole) was added slowly and the reaction mixture was brought to room temperature. After 16 h, the reaction mixture was quenched by adding methanol slowly and concentrated *in vacuo*. The resulting residue was mixed with water and

extracted with ethyl acetate (3×). The combined organic layers were dried over Na_2SO_4. After the volatiles were removed by rotary evaporation under reduced pressure, the crude product was purified using flash chromatography (silica gel; EtOAC/hexanes (1:9 → 1:4)) to yield **3-13** ($C_{51}H_{54}O_8$, 0.47 g, 81%) as syrup. TLC (EtOAc/hexanes (1:4)) R_f = 0.39. ^{1}H NMR ($CDCl_3$, 600 MHz) δ (ppm) 7.38–7.20 (30 H, m, Bn×6), 4.96 (1 H, J = 10.4 Hz, CH_2 of Bn), 4.82 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.81 (1 H, J = 10.4 Hz, CH_2 of Bn), 4.79 (1 H, J = 12.0 Hz, CH_2 of Bn), 4.72 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.67 (1 H, J = 12.0 Hz, CH_2 of Bn), 4.66 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.61 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.59 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.49 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.48 (2 H, J = 12.4 Hz, CH_2 of Bn), 4.43(1 H, J = 11.2 Hz, CH_2 of Bn), 4.13 (1 H, dd, $J_{5,6}$ = 1.2 Hz, $J_{4,5}$ = 10.0 Hz, H5), 3.98 (1 H, dd, $J_{2,3}$ = 9.6 Hz, $J_{3,4}$ = 9.2 Hz, H3), 3.88 (2 H, brs, H_{8a} and H_{8b}), 3.82 (1 H, dd, $J_{3,4}$ = 9.2 Hz, $J_{4,5}$ = 10.0 Hz, H4), 3.71 (1 H, dd, $J_{5,6}$ = 1.2 Hz, $J_{6,7}$ = 10.0 Hz, H6), 3.62 (1 H, ddd, $J_{6,7}$ = 10.0 Hz, $J_{7,8a}$ = 0.8 Hz, $J_{7,8b}$ = 2.4 Hz, H7), 3.47 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{2,3}$ = 9.6 Hz, H2), 3.34 (3 H, s, OMe). ^{13}C NMR ($CDCl_3$, 150 MHz) δ (ppm) 138.83, 138.62, 138.61, 138.48, 138.47, 138.29, 128.44, 128.36, 128.27, 128.26, 128.16, 128.01, 127.92, 127.86, 127.69, 127.57, 127.56, 127.55, 127.44, 127.43, 97.82, 82.83, 80.14, 78.76, 78.24, 78.16, 75.75, 74.73, 73.53, 73.29, 73.23, 72.49, 70.26, 69.58, 55.02. ESI-HRMS Calcd. for $C_{51}H_{54}O_8Na^+$ $[M + Na]^+$ 817.3711, found 837.3710.

1,8-Di-*O*-acetyl-2,3,4,6,7-penta-*O*-benzyl-α,β-D-*erythro*-D-*gluco*-octopyranose (3-14).

Compound **3-13** (1.98 g, 2.49 mmole) was dissolved in acetic acid (10 mL) and added acetic anhydride (10 mL). The reaction mixture was cooled to 0 °C in an ice-water bath and then added concentrated sulfuric acid (1 mL) dropwise. After stirring at 0 °C for 2 h, the reaction mixture was diluted with dichloromethane and quenched by adding saturated sodium bicarbonate solution slowly. The resulting solution was extracted with dichloromethane (3×). The combined organic lay-

ers were washed with water (1×) and brine (1×), and then dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation under reduced pressure, the crude product was purified using flash chromatography (silica gel; EtOAc/hexanes (1:9 → 1:2)) to yield **3-14** ($C_{47}H_{50}O_{10}$, 1.02 g, α:β = 6:1, 53%) as syrup. TLC (EtOAc/hexanes (1:2)) R_f = 0.55. α **anomer**: ^{1}H NMR ($CDCl_3$, 600 MHz) δ (ppm) 7.32–7.19 (25 H, m, Bn×5), 6.32 (1 H, d, $J_{1,2}$ = 3.6 Hz, H1), 4.96 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.89 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.82 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.72 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.70 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.63 (2 H, J = 10.8 Hz, CH_2 of Bn), 4.51 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.49 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.46 (1 H, dd, $J_{7,8a}$ = 2.4 Hz, $J_{8a,8b}$ = 12.0 Hz, H_{8a}), 4.28 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.24 (1 H, dd, $J_{4,5}$ = 8.8 Hz, $J_{5,6}$ = 9.6 Hz, H5), 4.04 (1 H, dd, $J_{7,8b}$ = 3.6 Hz, $J_{8a,8b}$ = 12.0 Hz, H_{8b}), 3.92 (1 H, dd, $J_{2,3}$ = 9.2 Hz, $J_{3,4}$ = 8.8 Hz, H3), 3.89–3.84 (2 H, m, H4 and H7), 3.80 (1 H, dd, $J_{5,6}$ = 9.6 Hz, $J_{6,7}$ = 10.0 Hz, H6), 3.57 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{2,3}$ = 9.2 Hz, H2), 2.05 (3 H, s, OAc), 1.94 (3 H, s, OAc). ^{13}C NMR ($CDCl_3$, 150 MHz) δ (ppm) 170.85, 169.47, 138.58, 138.25, 138.05, 137.96, 137.64, 128.47, 128.37, 128.32, 128.32, 128.28, 128.04, 128.00, 127.97, 127.89, 127.76, 127.66, 89.65, 82.31, 79.15, 78.00, 77.62, 76.75, 75.60, 75.04, 73.54, 73.19, 72.73, 72.22, 62.56, 20.97, 20.84.

β **anomer**: ^{1}H NMR ($CDCl_3$, 600 MHz) δ (ppm) 7.32–7.19 (25 H, m, Bn×5), 5.61 (1 H, $J_{1,2}$ = 8.0 Hz, H1), 4.89 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.84 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.82 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.78 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.77 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.61 (1 H, J = 12.0 Hz, CH_2 of Bn), 4.46–4.41 (3 H, m, CH_2 of Bn×2 and H_{8a}), 4.34–4.29 (2 H, CH_2 of Bn and H5), 4.06–4.02 (1 H, H_{8b}), 3.90–3.81 (3 H, H3, H4 and H6), 3.69 (1 H, ddd, $J_{6,7}$ = 8.8 Hz, $J_{7,8a}$ = $J_{7,8b}$ = 4.4 Hz, H7), 3.50 (1 H, dd, $J_{1,2}$ = 8.0 Hz, $J_{2,3}$ = 9.2 Hz, H2), 2.07 (3 H, s, OAc), 1.93 (3 H, s, OAc). ^{13}C NMR ($CDCl_3$, 150 MHz) δ (ppm) 206.86, 169.02, 138.31, 138.20, 138.17, 137.96, 137.94, 128.42, 128.39, 128.34, 128.30, 128.13, 128.02, 127.93, 127.79, 127.69,

127.56, 94.12, 85.24, 81.34, 78.16, 77.32, 76.46, 75.62, 75.45, 74.95, 74.80, 73.42, 72.20, 62.62, 30.88, 21.04, 20.82. ESI-HRMS Calcd. for $C_{47}H_{50}O_{10}Na^+$ $[M+Na]^+$ 797.3296, found 797.3309.

2,3,4,6,7-Penta-*O*-benzyl-α,β-D-*erythro*-D-*gluco*-octopyranose (3-15).

Compound **3-14** (0.93 g, 1.2 mmole) was dissolved in anhydrous methanol (40 mL) and added sodium methoxide methanol solution (0.5 M, 15 mL) dropwise. After stirring at ambient temperature for 1 h, the reaction mixture was neutralized by adding Dowex® 50W×8 (hydrogen form). The resins were removed by filtration, the solution was concentrated by rotary evaporation under reduced pressure, and the crude product **3-15** was used for later steps without further purification.

D-*Erythro*-D-*gluco*-octose α-1,8-bisphosphate (3-2).

Compound **3-15** (∼0.60 mg, 0.87 mmole) and tetrazole (0.45 M in acetonitrile, 12 mL, 5.4 mmole) was suspended in anhydrous toluene (10 mL) and the solution was evaporated in vacuo. The residue was then dissolved in anhydrous dichloromethane (60 mL) and cooled to −20 °C in a dry ice-acetone bath. Dibenzyl *N,N*-diisopropylphosphoramidite (1.7 mL, 5.06 mmole) was added dropwise over 15 min. The reaction solution was stirred at −20 °C for 1 h and then warmed to room temperature for another 1 h. The phosphorylation was complete as monitored by TLC. The reaction solution was cooled to 0 °C, added mCPBA (77%, 2.1 g, 9.4 mmole) portion-wise and allowed to stir at room temperature for 1 h. The oxidation was complete judging by TLC. The reaction mixture was diluted by adding dichloromethane and quenched by adding saturated $Na_2S_2O_3$ solution. The resulting solution was extracted with dichloromethane (3×). The combined organic layers were washed with water (1×) and then dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation under reduced pressure, the crude product was purified using flash chromatography (silica gel; EtOAC/hexanes (1:2)). The eluted fractions containing product were concentrated

and further purified using HPLC on an Econosil C18 semiprep column, 10 μm, 250 × 10 mm (Alltech) with detection at λ = 254 nm. Mobile phase A was water, while mobile phase B was acetonitrile. The separation was obtained at a flow rate of 4 mL/min with a gradient program that allowed for 8 min at 80% B followed by a 2-min step that raised eluent B to 90%. Then washing at 90% B for 10 min and re-equilibration at 80% B was performed in a total time of 30 min. The retention time of the product 2,3,4,6,7-penta-*O*-benzyl-α,β-D-*erythro*-D-*gluco*-octopyranose 1,8-bis-[bis(benzyl) phosphate] (**3-16**) was 17 min ($C_{71}H_{72}O_{14}P_2$, 0.11 g, 10%). ^{1}H NMR ($CDCl_3$, 500 MHz) δ (ppm) 7.36–7.17 (45 H, m, Bn×9), 5.93 (1 H, dd, $J_{1,2}$ = 3.0 Hz, $J_{1,P}$ = 7.5 Hz, H1), 5.02–4.84 (8 H, m, CH_2 of P(O)$(OBn)_2$×8), 4.91 (1 H, J = 11.0 Hz, CH_2 of Bn), 4.84 (1 H, J = 10.5 Hz, CH_2 of Bn), 4.79 (1 H, J = 11.0 Hz, CH_2 of Bn), 4.76 (1 H, J = 11.0 Hz, CH_2 of Bn), 4.63 (2 H, J = 11.5 Hz, CH_2 of Bn×2), 4.57 (1 H, J = 11.0 Hz, CH_2 of Bn), 4.49 (1 H, J= 11.5 Hz, CH_2 of Bn), 4.45 (1 H, J = 11.5 Hz, CH_2 of Bn), 4.38 (1 H, d, $J_{4,5}$ = 10.0 Hz, H5), 4.34 (1 H, ddd, $J_{7,8a}$ = 1.5 Hz, $J_{8a,8b}$ = 11.0 Hz, $J_{8a,P}$ = 6.0 Hz, H_{8a}), 4.27 (1 H, J = 11.5 Hz, CH_2 of Bn), 4.11 (1 H, ddd, $J_{7,8b}$ = 3.5 Hz, $J_{8a,8b}$ = 11.0 Hz, $J_{8b,P}$ = 6.0 Hz, H_{8b}), 3.91 (1 H, dd, $J_{2,3}$ = 9.5 Hz, $J_{3,4}$ = 9.5 Hz, H3), 3.87–3.83 (3 H, m, H4, H6 and H7), 3.48 (1 H, ddd, $J_{1,2}$ = 3.0 Hz, $J_{2,3}$ = 9.5 Hz, $J_{2,P}$ = 3.0 Hz, H2). ^{13}C NMR ($CDCl_3$, 125 MHz) δ (ppm) 138.48, 138.13, 137.88, 137.79, 137.61, 135.89 (d, $J_{C-OBn,P}$ = 7.3 Hz), 135.81 (d, $J_{C,P}$ = 6.9 Hz), 135.78 (d, $J_{C,P}$ = 7.3 Hz), 135.70 (d, $J_{C,P}$ = 6.9 Hz), 128.49, 128.46, 128.38, 128.36, 128.30, 128.29, 128.20, 128.07, 127.82, 127.77, 127.73, 127.63, 127.58, 95.22 (d, $J_{C1,P}$ = 5.9 Hz), 81.57, 79.45 (d, $J_{C2,P}$ = 7.4 Hz), 77.80, 77.15, 75.52, 74.95, 73.55, 72.97, 72.67, 72.32, 69.13 (d, $J_{C-OBn,P}$ = 3.8 Hz), 69.10 (d, $J_{C-OBn,P}$ = 5.5 Hz), 66.41 (d, $J_{C8,P}$ = 5.9 Hz). ESI-HRMS Calcd. for $C_{71}H_{72}O_{14}P_2Na^+$ $[M + Na]^+$ 1233.4290, found 1233.4290.

The purified **3-16** (110 mg, 0.09 mmole) and 5% Pd/C (50 mg) in a methanol solution (25 mL) was stirred under hydrogen atmosphere (1 atm) for 1 h. The reaction

mixture was filtered through a pad of Celite and the filtrate after concentration yield bisphosphate **3-2** ($C_8H_{18}O_{14}P_2$, 10 mg, 27%) as white solids. ^{1}H NMR (D_2O, 600 MHz) δ (ppm) 5.35 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{1,P}$ = 6.6 Hz, H1), 3.99 (1 H, d, $J_{4,5}$ = 9.0 Hz, H5), 3.92–3.82 (4 H, m, H6, H7, H8$_a$ and H8$_b$), 3.64 (1 H, dd, $J_{2,3}$ = 9.0 Hz, $J_{3,4}$ = 9.0 Hz, H3), 3.60 (1 H, dd, $J_{3,4}$ = 9.0 Hz, $J_{4,5}$ = 9.0 Hz, H4), 3.41 (1 H, brd, $J_{2,3}$ = 9.0 Hz, H2). ^{13}C NMR (D_2O, 125 MHz) δ (ppm) 94.81, 72.87, 72.26, 71.11, 70.72, 70.39, 68.89, 66.72. ^{31}P NMR (D_2O, 243 MHz) δ (ppm) 2.53, 0.49. ESI-HRMS Calcd. for $C_8H_{18}O_{14}P_2Na^+$ $[M + Na]^+$ 423.0064, found 423.0063.

3.2.8 Synthesis of D-*erythro*-D-*gluco*-octose α-1-phosphate (3-3)

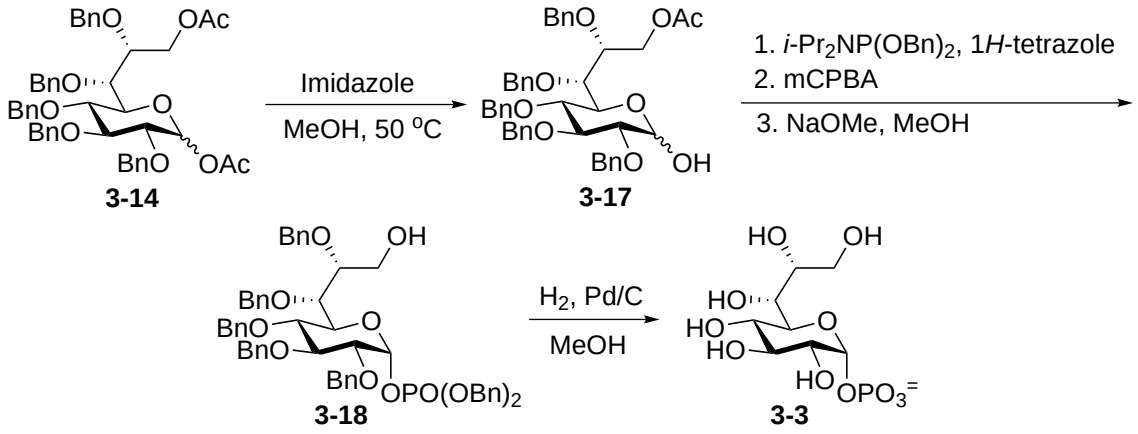


Figure 3.3: Synthesis of octose 1-phosphate (**3-3**).

Methyl 8-*O*-Acetyl-2,3,4,6,7-penta-*O*-benzyl-α,β-D-*erythro*-D-*gluco*-octopyrano -side (3-17).

The synthesis of **3-3** was shown in Figure 3.3. Compound **3-14** (0.70 g, 0.90 mmole) was dissolved in methanol (50 mL) and added imidazole methanolic solution (1 M, 50 mL). The reaction mixture was heated to 40 °C in an oil bath. After stirring at 40 °C for 20 h, the reaction mixture was concentrated under reduced pressure. The residue was extracted with EtOAc (3×). The combined organic layers were dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation

under reduced pressure, the crude product was purified using flash chromatography (silica gel; EtOAC/hexanes (1:4 → 1:2)) to yield **3-17** ($C_{45}H_{48}O_9$, 0.40 g, α:β = 3:2, 60%) as syrup. TLC (EtOAc/hexanes (1:2)) R_f = 0.47. **α anomer**: ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) 7.38–7.19 (25 H, m, Bn×5), 5.23 (1 H, dd, $J_{1,2}$ = 4.0 Hz, H1), 4.95 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.84 (2 H, J = 10.8 Hz, CH_2 of Bn×2), 4.77 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.73 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.71 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.61 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.51 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.49 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.47 (1 H, dd, $J_{7,8a}$ = 2.0 Hz, $J_{8a,8b}$ =11.6 Hz, H_{8a}), 4.39 (1 H, d, $J_{4,5}$ = 10.4 Hz, H5), 4.36 (1 H, J = 11.2 Hz, CH_2 of Bn), 4.06 (1 H, dd, $J_{7,8b}$ = 3.2 Hz, $J_{8a,8b}$ =11.6 Hz, H_{8b}), 3.98 (1 H, dd, $J_{2,3}$ = 9.6 Hz, $J_{3,4}$ = 9.2 Hz, H3), 3.90–3.86 (1 H, m, H7), 3.84 (1 H, d, $J_{6,7}$ = 9.2 Hz, H6), 3.81 (1 H, dd, $J_{3,4}$ = 9.2 Hz, $J_{4,5}$ = 9.2 Hz, H4), 3.49 (1 H, dd, $J_{1,2}$ = 4.0 Hz, $J_{2,3}$ = 9.6 Hz, H2), 1.92 (3 H, s, OAc). ^{13}C NMR ($CDCl_3$, 125 MHz) δ (ppm) 170.89, 138.68, 138.45, 137.95, 137.91, 137.90, 128.5, 128.41, 128.36, 128.31, 128.29, 128.24, 128.20, 128.15, 128.09, 127.94, 127.91, 127.87, 127.81, 127.75, 127.68, 127.61, 127.57, 127.55, 91.10, 82.25, 80.23, 78.64, 77.10, 76.63, 75.57, 74.74, 72.29, 70.28, 62.86, 20.82.

β **anomer**: ^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm) 7.38-7.19 (25 H, m, Bn×5), 4.96 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.93 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.84 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.81 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.77 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.71 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.71 (1 H, $J_{1,2}$ = 9.2 Hz, H1), 4.61 (1 H, J = 10.8 Hz, CH_2 of Bn), 4.54 (1 H, J = 11.6 Hz, CH_2 of Bn), 4.52–4.45 (3 H, m, CH_2 of Bn×2 and H_{8a}), 4.39 (1 H, d, $J_{4,5}$ = 10.4 Hz, H5), 4.06 (1 H, dd, $J_{7,8b}$ = 3.2 Hz, $J_{8a,8b}$ =11.6 Hz, H_{8b}), 3.90–3.86 (1 H, m, H7), 3.85–3.77 (2 H, m, H6 and H4), 3.63 (1 H, dd, $J_{2,3}$ = 8.8 Hz, $J_{3,4}$ = 8.8 Hz, H3), 3.35 (1 H, dd, $J_{1,2}$ = 9.2 Hz, $J_{2,3}$ = 8.8 Hz, H2), 1.93 (3 H, s, OAc). ^{13}C NMR ($CDCl_3$, 125 MHz) δ (ppm) 170.89, 138.52, 138.37, 138.34, 137.89, 137.78, 128.5, 128.41, 128.36, 128.31, 128.29, 128.24, 128.20, 128.15, 128.09, 127.94, 127.91, 127.87, 127.81, 127.75, 127.68, 127.61, 127.57, 127.55,

97.80, 85.04, 83.23, 78.45, 76.57, 75.57, 74.80, 62.78, 19.98. ESI-HRMS Calcd. for $C_{45}H_{48}O_9Na^+$ $[M + Na]^+$ 755.3191, found 755.3195.

D-*Erythro*-D-*gluco*-octose α-1-phosphate (3-3).

Compound **3-17** (0.20 g, 0.27 mmole) and tetrazole (0.45 M in acetonitrile, 2.7 mL, 1.2 mmole) was suspended in anhydrous toluene (10 mL) and the solution was evaporated in in vacuo. The residue was then dissolved in anhydrous dichloromethane (10 mL) and cooled to $-20\,^{\circ}C$ in a dry ice-acetone bath. Dibenzyl *N,N*-diisopropyl-phosphoramidite (0.35 mL, 1.06 mmole) was added dropwise over 15 min. The reaction solution was stirred at $-20\,^{\circ}C$ for 1 h and then warmed to room temperature for another 1 h. The phosphorylation was complete by checking TLC. The reaction solution was cooled to $0\,^{\circ}C$, added mCPBA (77%, 0.4 g, 1.79 mmole) portion-wise and stirred at room temperature for 1 h. The oxidation was complete by checking TLC. The reaction mixture was diluted by adding dichloromethane and quenched by adding saturated $Na_2S_2O_3$ solution. The resulting solution was extracted with dichloromethane ($3\times$). The combined organic layers were washed with water ($1\times$) and then dried over anhydrous Na_2SO_4. After the volatiles were removed by rotary evaporation under reduced pressure, the crude product was dissolved in methanol (5 mL) and added sodium methoxide methanol solution (0.5 M, 1 mL). After stirring at ambient temperature for 4 h, the reaction mixture was neutralized by adding Dowex® 50W$\times$8 (hydrogen form). The resins were removed by filtration and the crude product was purified using flash chromatography (silica gel; EtOAC/hexanes (1:2)) to yield 2,3,4,6,7-penta-*O*-benzyl-α-D-*erythro*-D-*gluco*-octopyranose 1-[bis(benzyl) phosphate]. (**3-18**, $C_{57}H_{59}O_{11}P$, 60 mg, 23%) as syrup. TLC (EtOAc/hexanes (1:1)) R_f = 0.61. ^{1}H NMR ($CDCl_3$, 500 MHz) δ (ppm) 7.37–7.14 (35 H, m, Bn$\times$7), 5.98 (1 H, dd, $J_{1,2}$ = 3.0 Hz, $J_{1,P}$ = 7.0 Hz, H1), 5.04 (1 H, dd, J = 11.5 Hz, $J_{H,P}$ = 7.0 Hz, CH_2 of P(O)$(OBn)_2$), 5.03 (1 H, dd, J = 12.0 Hz, $J_{H,P}$ = 7.0 Hz, CH_2 of P(O)$(OBn)_2$), 5.00 (1 H, dd, J = 12.0 Hz, $J_{H,P}$ = 7.0 Hz, CH_2 of P(O)$(OBn)_2$),

4.94 (1 H, J = 11.0 Hz, CH_2 of Bn), 4.93 (1 H, dd, J = 11.5 Hz, $J_{H,P}$ = 7.0 Hz, CH_2 of $P(O)(OBn)_2$), 4.89 (1 H, J = 10.5 Hz, CH_2 of Bn), 4.81 (1 H, J = 11.5 Hz, CH_2 of Bn), 4.79 (1 H, J = 11.0 Hz, CH_2 of Bn), 4.68 (1 H, J = 11.5 Hz, CH_2 of Bn), 4.66 (1 H, J = 11.5 Hz, CH_2 of Bn), 4.63 (1 H, J= 10.5 Hz, CH_2 of Bn), 4.52 (1 H, J = 11.5 Hz, CH_2 of Bn), 4.37 (1 H, J = 11.5 Hz, CH_2 of Bn), 4.36 (1 H, d, $J_{4,5}$ = 11.0 Hz, H5), 4.34 (1 H, J = 11.5 Hz, CH_2 of Bn), 3.94 (1 H, dd, $J_{2,3}$ = 9.5 Hz, $J_{3,4}$ = 9.0 Hz, H3), 3.85 (1 H, dd, $J_{3,4}$ = 9.0 Hz, $J_{4,5}$ = 11.0 Hz, H4), 3.81 (1 H, d, $J_{6,7}$ = 9.5 Hz, H6), 3.72 (1 H, ddd, $J_{6,7}$ = 9.5 Hz, $J_{7,8ab}$ = 3.5 Hz, H7), 3.63 (2 H, brd, $J_{7,8ab}$ = 3.5 Hz, H_{8a} and H_{8b}), 3.52 (1 H, ddd, $J_{1,2}$ = 3.0 Hz, $J_{2,3}$ = 9.5 Hz, $J_{2,P}$ = 3.0 Hz, H2). ^{13}C NMR ($CDCl_3$, 125 MHz) δ (ppm) 138.45, 138.17, 137.89, 137.86, 137.60, 135.88 (d, $J_{C-OBn,P}$ = 8.0 Hz), 135.77 (d, $J_{C,P}$ = 7.5 Hz), 128.39, 128.37, 128.35, 128.29, 128.22, 128.10, 128.05, 127.93, 127.92, 127.85, 127.82, 127.75, 127.71, 127.64, 127.62, 127.60, 95.19 (d, $J_{C1,P}$ = 5.9 Hz), 81.69, 79.48 (d, $J_{C2,P}$ = 6.9 Hz), 78.38, 78.36, 77.58, 75.62, 74.83, 73.70, 72.93, 72.78, 72.08, 69.14 (d, $J_{C-OBn,P}$ = 4.9 Hz), 69.05 (d, $J_{C-OBn,P}$ = 5.5 Hz). ^{31}P ($CDCl_3$, 202 MHz) δ (ppm) 1.18. ESI-HRMS Calcd. for $C_{57}H_{59}O_{11}PNa^+$ $[M + Na]^+$ 973.3687, found 973.3686.

The purified **3-18** (∼60 mg, 0.08 mmole) was dissolved in methanol and added 5% Pd/C (5 mg). The reaction mixture was stirred under hydrogen atmosphere (1 atm) for 1 h. The reaction solution was filtered through a pad of Celite and the filtrate after concentration yield monophosphate **10** ($C_8H_{17}O_{11}P$, 26 mg, quantitative yield) as white solids. ^1H NMR (D_2O, 400 MHz) δ (ppm) 5.31 (1 H, dd, $J_{1,2}$ = 3.2 Hz, $J_{1,P}$ = 6.8 Hz, H1), 3.95 (1 H, dd, $J_{4,5}$ = 10.0 Hz, $J_{5,6}$ = 2.4 Hz, H5), 3.80 (1 H, ddd, $J_{6,7}$ = 7.2 Hz, $J_{7,8a}$ = 3.2 Hz, $J_{7,8b}$ = 6.4 Hz, H7), 3.75 (1 H, dd, $J_{4,5}$ = 2.4 Hz, $J_{6,7}$ = 7.2 Hz, H6), 3.67 (1 H, dd, $J_{7,8a}$ = 3.2 Hz, $J_{8a,8b}$ = 12.0 Hz, H_{8a}), 3.62 (1 H, dd, $J_{2,3}$ = 9.6 Hz, $J_{3,4}$ = 9.6 Hz, H3), 3.53 (1 H, dd, $J_{3,4}$ = 9.6 Hz, $J_{4,5}$ = 9.6 Hz, H4), 3.49 (1 H, dd, $J_{7,8b}$ = 6.4 Hz, $J_{8a,8b}$ = 12.0 Hz, H_{8b}), 3.36 (1 H, brd, $J_{2,3}$ = 9.6 Hz, H2). ^{13}C NMR (D_2O, 125 MHz) δ (ppm) 93.49, 73.50, 72.02, 72.02, 71.57, 71.57, 70.61, 63.21. ^{31}P (D_2O, 202

MHz) δ (ppm) 2.13. ESI-HRMS Calcd. for $C_8H_{16}O_{11}P^-$ $[M-H]^-$ 319.0436, found 319.0436.

3.2.9 Synthesis of GDP-D-*erythro*-D-octose (3-5)

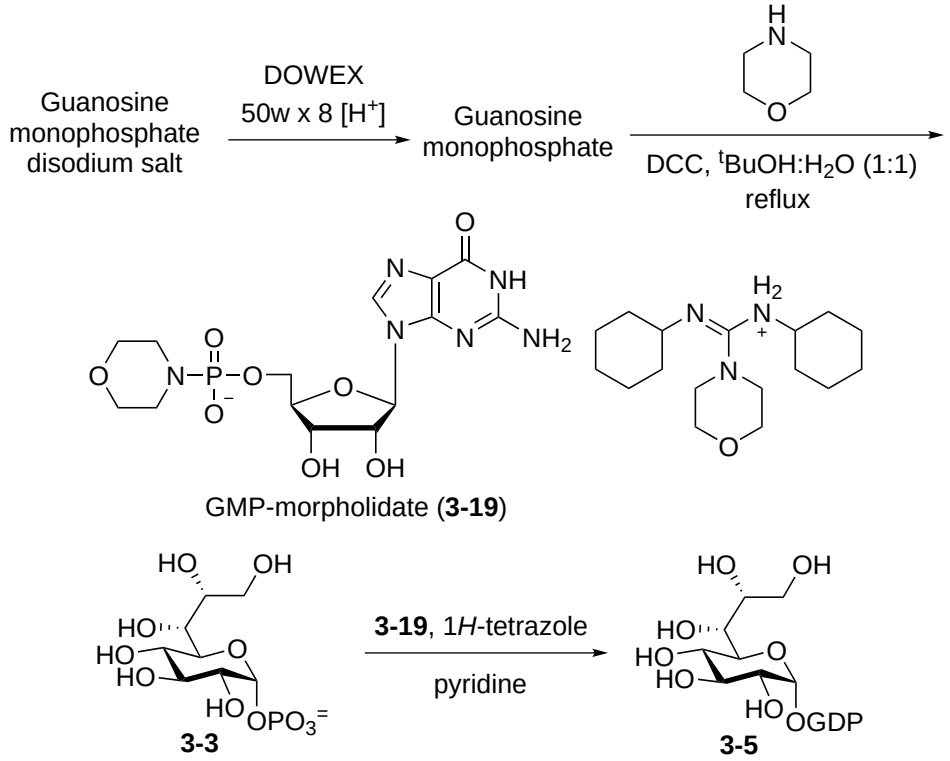


Figure 3.4: Synthesis of octose 1,8-bisphosphate (**3-5**)

Guanosine-5′-monophsophate morpholidate (3-19)

The overall synthetic scheme of compound **3-5** is shown in Figure 3.4. Guanosine monophosphate disodium salt (4.10 g, 10.07 mmole) in water (50 mL) was applied to a column packed with Dowex® 50W×8 (hydrogen form). The eluted fractions containing GMP concentrated by rotary evaporation and then lyophilized overnight. The GMP were dissolved in 100 mL of ^{t}BuOH:H_2O (1:1) co-solvent and added morpholine (2.4 mL, 27.5 mmole). The reaction mixture was heated to reflux for 3 h, during which time a solution of *N*,*N*′-dicyclohexylcarbodiimide (6.2 g, 30.0 mmol) in ^{t}BuOH (15 mL) was added dropwise. After refluxing the reac-

tion solution for an additional 2 h, the reaction mixture was cooled down to room temperature. The precipitates were removed by filtration and the resulting solution was concentrated by rotary evaporation. The residues were then re-dissolved in water and the insoluble precipitates were removed by filtration. The solution was then freeze dried. The 4-morpholine-*N*,*N*′-dicylcohexylcarboxamidine salt of **3-19** was obtained as a colorless powder (3.88 g, 77%). ^{1}H NMR (D_2O, 400 MHz) δ (ppm) 7.92 (1 H, s, guanine), 5.79 (1 H, d, $J_{1,2}$ = 5.2 Hz, H1), 4.38 (1 h, dd, $J_{2,3}$ = $J_{3,4}$ = 4.8 Hz, H3), 4.16 (1 H, brs, H4), 3.91–3.84 (2 H, m, H5), 3.58 (4 H, t, J = 4.8, $CH_2\times2$ of morpholine of counter ion), 3.43 (4 H, t, J = 4.4, $CH_2\times2$ of morpholine), 3.31 (4 H, t, J = 4.8, $CH_2\times2$ of morpholine of counter ion), 3.21–3.16 (2 H, m, CH×2 of cyclohexyl of counter ion), 2.84–2.76 (4 H, m, $CH_2\times2$ of morpholine), 1.79–0.98 (20 H, m, $CH_2\times10$ of cyclohexyl of counter ion). ^{31}P NMR (D_2O, 162 MHz) δ (ppm) 7.40. Analytical data were consistent with the commercial standards.[144]

GDP-D-*erythro*-D-octose (3-5).

Compound **3-3** (6.4 mg, 0.02 mmol) was dissolved in water (5 mL) and added Dowex® 50W×8 (tri-*n*-butylammonium salt form). The reaction mixture was stirred at 4 °C for 16 h and then filtered to remove the resins. The solution was lyophilized to afford the tri-*n*-butylammonium salt of **3-3**. To the tri-*n*-butylammonium salt of **3-3** in anhydrous pyridine (10 mL) were added 1*H*-tetrazole (0.45 M in acetonitrile, 0.2 mL, 0.09 mmol) and GMP-morpholidate (**3-19**, 30 mg, 0.03 mmole). After stirring for 3 days at room temperature, the reaction mixture was concentrated *in vacuo*. The residue was dissolved in water (10 mL) and then washed with diethyl ether (3×). The aqueous solution was concentrated and injected to a Dionex CarboPacPA1 semiprep column, 250 × 9 mm with detector set at λ = 254 nm. Mobile phase A was water, while mobile phase B was 1 M NH_4HCO_3. The separation was obtained at a flow rate of 3 mL/min with a gradient program: 0-2 min 30% B, 2-12 min 30-40% B, 12-20 min 40% B, 20-21 min 40-70% B, 21-31 min 70%

B, 31-32 min 70-30% B, 32-40 min 30% B. The retention time of the product **3-5** *N,N*-dicyclohexyl-4-morpholinecarboxamidine salt was 15 min ($C_{35}H_{60}N_8O_{19}P_2$, 3.7 mg, 19%). ^{1}H NMR (D_2O, 500 MHz) δ (ppm) 7.89 (1 H, s, guanine), 5.72 (1 H, d, $J_{1,2}$ = 6.0 Hz, ribose-H1), 5.29 (1 h, dd, $J_{1,2}$ = 3.5 Hz, $J_{1,P}$ = 7.0 Hz, octose-H1), 4.54 (1 H, dd, $J_{1,2}$ = 6.0 Hz, $J_{2,3}$ = 5.5 Hz ribose-H2), 4.29 (1 H, brt, $J_{2,3}$ = 5.5 Hz, ribose-H3), 4.15–4.14 (1 H, m, ribose-H4), 4.00 (2 H, brs, ribose-H5), 3.85 (1 H, brd, $J_{4,5}$ = 9.5 Hz, octose-H5), 3.82–3.68 (2 H, m, octose-H6 and H7), 3.60 (1 H, brd, $J_{8a,8b}$ = 12.0 Hz, octose-H8$_a$), 3.59–3.57 (4 H, m, morpholine), 3.37 (1 H, dd, $J_{2,3}$ = $J_{3,4}$ = 9.5 Hz, octose-H3), 3.35 (1 H, dd, $J_{3,4}$ = $J_{4,5}$ = 9.5 Hz, octose-H4), 3.43 (1 H, dd, $J_{7,8b}$ = 6.5 Hz, $J_{8a,8b}$ = 12.0 Hz, octose-H8$_b$), 3.34–3.18 (1 H, m, octose-H2), 3.23 (4 H, brs, morpholine), 3.13–3.11 (2 H, m, C×H2 of cyclohexyl), 1.70–0.92 (20 H, m, CH_2×10 of cyclohexyl). ^{31}P NMR (D_2O, 202 MHz) δ (ppm) -10.42 (J = 20.0 Hz), -12.10 (J = 20.0 Hz). ESI-HRMS Calcd. for $C_{18}H_{28}N_5O_{18}P_2^-$ $[M-H]^-$ 664.0910, found 664.0922.

3.2.10 LmbK and LmbO Reaction Assays

LmbK reaction using the synthetic bisphosphate 3-2 as substrate

The *N*-His$_6$-LmbK protein (44 μM final concentration) was incubated with D-*erythro*-D-*gluco*-octose α-1,8-bisphosphate (**3-2**, 3 mM) and $MgCl_2$ (10 mM) in 25 mM Tris buffer (pH 8.0) at ambient temperature for 2.5 h or 5 h. A control reaction omitting the enzyme was similarly prepared. At the end of the incubation, each reaction mixture was filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the protein. HPLC analysis of the filtrate was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm). The samples were eluted by using water as mobile phase A and 1 M NH_4OAc as mobile phase B. The separation was obtained at a flow rate of 1 mL/min with a gradient program: 0-0.5 min 5% B, 0.5-1.5 min 5-50% B, 1.5-10 min 50% B, 10-11 min 50-5% B, 11-20 min 5% B. The elution was monitored by a Corona CAD detector.

NMR characterization of LmbK reaction using the synthetic bisphosphate 3-2 as substrate

The reaction solution (0.8 mL $\times$ 3) contained compound **3-2** (12.5 mM) and *N*-His_6-LmbK (18.8 M) in 50 mM $NH_4 \cdot HCO_3$ buffer (pH 8) containing 10 mM $MgCl_2$. Prior to incubation, the *N*-His_6-LmbK protein stock solution was dialyzed against $NH_4 \cdot HCO_3$ buffer (pH 8) containing 10 mM $MgCl_2$ to remove excess glycerol. The reaction mixture was incubated at ambient temperature for 16 h and filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the protein. The filtrate was then lyophilized and redissolved in D_2O for the NMR analysis.

LmbK and LmbO reaction using the synthetic bisphosphate 3-2 and various nucleotide triphosphates as substrates

The purified *N*-His_6-LmbK protein (10 μM final concentration) and the refolded *N*-His_6-LmbO (19 μM final concentration) were incubated with bisphosphate **3-2** (2 mM), NTP (3 mM), inorganic pyrophosphatase (1 unit) and $MgCl_2$ (10 mM) in 25 mM Tris buffer (pH 8.0) at ambient temperature for 8 h. At the end of the incubation, each reaction mixture was filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the protein. HPLC analysis of the filtrate was performed using a Dionex CarboPac PA1 analytical column (4 $\times$ 250 mm). Mobile phase A was water and mobile phase B was 1 M NH_4OAc. The separation was achieved at a flow rate of 1 mL/min. The samples containing CTP, TTP and UTP were analyzed by gradient method (1): 0-2 min 30% B, 2-10 min 30-90% B, 10-20 min 90% B, 20-22 min 90-30% B, 22-30 min 30% B. The samples containing ATP and GTP were analyzed by gradient method (2): 0-2 min 50% B, 2-10 min 50-90% B, 10-20 min 90% B, 20-22 min 90-50% B, 22-30 min 50% B. The detector wavelength was set at 267 nm for cytosine and thymine, 262 nm for uracil, 259 nm for adenine, and 254 nm for guanine. The HPLC fraction containing the product

(retention time ~9 min) was collected. The collected fraction was lyophilized and redissolved in deionized water prior to ESI-MS analysis (negative ion detection mode).

LmbO reaction using the synthetic monophosphate 3-3 and various nucleotide triphosphates as substrates

The refolded *N*-His_6-LmbO protein (19 μM final concentration) was incubated with monophosphate **3-3** (2 mM), NTP (3 mM), inorganic pyrophosphatase (1 unit) and $MgCl_2$ (10 mM) in 25 mM Tris buffer (pH 8.0) at ambient temperature for 8 h. At the end of the incubation, each reaction mixture was filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the protein. HPLC analysis of the filtrate was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm) with the abovementioned methods.

LmbK and LmbO reaction using GTP and the synthetic bisphosphate 3-2 or monophosphate 3-3 as substrates

The purified *N*-His_6-LmbK protein (10 μM final concentration) and the refolded *N*-His_6-LmbO protein (19 μM final concentration) was incubated with bisphosphate **3-2** (5 mM), GTP (2 mM), inorganic pyrophosphatase (1 unit) and $MgCl_2$ (10 mM) in 25 mM Tris buffer (pH 8.0) at the ambient temperature for 18 h. A control reaction omitting LmbK and another reaction containing the refolded *N*-His_6-LmbO protein (19 μM final concentration) and monophosphate **3-3** (5 mM) were similarly prepared. At the end of incubation, each reaction mixture was filtered through a YM-10 membrane to remove the protein. HPLC analysis of the filtrate was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm) with the method: mobile phase A was water and mobile phase B was 1 M NH_4OAc. The separation was obtained at a flow rate of 1 mL/min with a gradient program: 0-2 min 60% B, 2-14 min 60-90% B, 14-24 min 90% B, 24-26 min 90-60%

B, 26-34 min 60% B. The detector wavelength was set at 254 nm.

NMR characterization of LmbO product

The reaction solution (0.5 mL × 6) containing *N*-His_6-LmbO protein (19 μM final concentration), monophoshate **3-3** (5 mM), GTP (8 mM), inorganic pyrophosphatase (1 unit) and $MgCl_2$ (10 mM) in 25 mM Tris buffer (pH 8.0) was incubated at the ambient temperature for 16 h. The reaction solution was filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove the protein. The filtrate was then purified using HPLC using a Dionex CarboPac PA1 semi-prep column (9 × 250 mm) with the method: mobile phase A was water and mobile phase B was 1 M NH_4OAc. The separation was obtained at a flow rate of 3 mL/min with a gradient program: 0-2 min 60% B, 2-14 min 60-90% B, 14-24 min 90% B, 24-26 min 90-60% B, 26-34 min 60% B. The detector wavelength was set at 254 nm. The retention time of the LmbO-product was 20 min. GDP-D-*erythro*-α-D-*gluco*-octose (**3-5**). ^{1}H NMR (D_2O, 600 MHz) δ (ppm) 8.09 (1 H, s, guanine), 5.92 (1 H, d, $J_{1,2}$ = 6.0 Hz, ribose-H1), 5.57 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{1,P}$ = 7.2 Hz, octose-H1), 4.49 (1 H, brt, $J_{2,3}$ = 3.6 Hz, ribose-H3), 4.23 (1 H, brs, ribose-H4), 4.20 (2 H, brs, ribose-H5), 4.06 (1 H, dd, $J_{4,5}$ = 9.6 Hz, $J_{5,6}$ = 3.0 Hz, octose-H5), 3.91 (1 H, ddd, $J_{6,7}$ = 7.8 Hz, $J_{7,8a}$ = 3.0 Hz, $J_{7,8b}$ = 6.6 Hz, octose-H7), 3.88 (1 H, $J_{5,6}$ = 3.0 Hz, $J_{6,7}$ = 7.8 Hz, octose-H6), 3.78 (1 H, dd, $J_{7,8a}$ = 3.0 Hz, $J_{8a,8b}$ = 12.0 Hz, octose-H8$_a$), 3.71 (1 H, dd, $J_{2,3}$ = $J_{3,4}$ = 9.6 Hz, octose-H3), 3.68 (1 H, dd, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, octose-H4), 3.58 (1 H, dd, $J_{7,8b}$ = 6.0 Hz, $J_{8a,8b}$ = 12.0 Hz, octose-H8$_b$), 3.55-3.50 (1 H, m, octose-H2). ^{31}P NMR (D_2O, 243 MHz) δ (ppm) −11.16 (19.9 Hz), −12.85 (19.9 Hz). ESI-HRMS Calcd. for $C_{18}H_{28}N_5O_{18}P_2[M-Na]^-$ 664.0910; Found: 664.6922.

LmbK and LmbO reaction using GTP and various glucose phosphates as substrates

The purified *N*-His_6-LmbK protein (10 μM final concentration) and the refolded *N*-His_6-LmbO protein (19 μM final concentration) was incubated with D-glucose α-1,6-bisphosphate (2 mM), GTP (2 mM), inorganic pyrophosphatase (1 unit) and $MgCl_2$ (10 mM) in 25 mM Tris buffer (pH 8.0) at the ambient temperature for 6 h. Other reactions containing the refolded *N*-His_6-LmbO protein only (19 μM final concentration) and D-glucose α-1-phosphate (2 mM) or D-glucose β-1-phosphate (2 mM) were similarly prepared. Each reaction mixture was filtered through a YM-10 membrane to remove the protein. HPLC analysis of the filtrate was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm) with the method: mobile phase A was water and mobile phase B was 1 M NH_4OAc. The separation was obtained at a flow rate of 1 mL/min with a gradient program: 0-2 min 50% B, 2-10 min 50-90% B, 10-20 min 90% B, 20-22 min 90-50% B, 22-30 min 50% B. The detector wavelength was set at 254 nm.

LmbO reaction using GTP and various sugar 1-phosphates as substrates

The refolded *N*-His_6-LmbO protein (19 μM final concentration) was incubated with different sugar 1-phosphates (5 mM), GTP (2 mM), inorganic pyrophosphatase (1 unit) and $MgCl_2$ (10 mM) in 25 mM Tris buffer (pH 8.0) at the ambient temperature for 16 h. Sugar substrates includes D-glucose α-1-phosphate, D-galactose α-1-phosphate, D-mannose α-1-phosphate, and compound **10**. Each reaction mixture was filtered through a YM-10 membrane to remove the protein. HPLC analysis of the filtrate was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm) with the method: mobile phase A was water and mobile phase B was 1 M NH_4OAc. The separation was obtained at a flow rate of 1 mL/min with a gradient program: 0-2 min 10% B, 2-10 min 10-50% B, 10-25 min

50-90% B, 25-30 min 90%B, 32-32 min 90-10% B, 32-40 min 10% B. The detector wavelength was set at 254 nm.

3.3 Results and Discussion

3.3.1 Overexpression and Purification of LmbP, LmbK and LmbO in *E. coli*

To verify the proposed pathways shown in Figure 3.1, the *lmbP*, *lmbK* and *lmbO* genes were amplified from the genomic DNA of *S. lincolnensis* ATCC 25466 and individually cloned into pET24b(+) or pET28b(+) vectors. The recombinant LmbP, LmbK and LmbO with *C*- or *N*-terminal His_6 tags were overexpressed in *E. coli*. LmbK could be purified to near homogeneity as soluble protein (Figure 3.5), but both LmbP and LmbO were expressed only as inclusion bodies. Since CcbP and CcbO encoded in celesticetin biosynthetic were predicted to be functionally identical to LmbP and LmbO, respectively, plasmids containing *ccbP* and *ccbO* genes were constructed to pursue the soluble proteins. Despite the efforts, both the overexpressed CcbP and CcbO proteins remained in the inclusion bodies. The denatured LmbO and CcbO could be solubilized from the inclusion bodies and refolded into soluble form through an optimized refolding protocol. However, attempts to obtain the soluble form of LmbP or CcbP using a similar approach were unsuccessful.

3.3.2 Functional Characterization of LmbK

The formation of intermediate **3-2** was proposed as the first reaction in NDP-octose biosynthesis. Accordingly, the availability of intermediate **3-2** is not only important for the verification of the proposed pathway, but also crucial for the examination of which route is favored for the later steps. The lack of functional LmbP, however, hampered the investigation of lincomycin A biosynthetis. To overcome this obstacle, the putative substrate for the subsequent enzymes (LmbK in

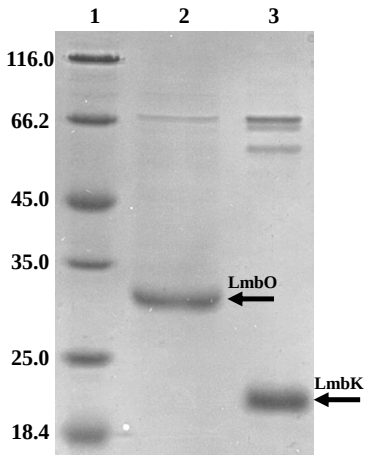


Figure 3.5: SDS-PAGE gel of the purified *N*-His_6-tagged LmbK (22.5 kDa, lane 3) and the refolded *N*-His_6-tagged LmbO (26.6 kDa, lane 2). The molecular weight markers are shown in lane 1. Minor impurities are possibly resulted from non-specific binding of other proteins with Ni-NTA resins.

route A and LmbO in route B) were chemically synthesized (Figure 3.2).

The synthetic bisphosphate **3-2** was first incubated with LmbK and the reaction mixtures were analyzed by HPLC equipped with a Corona charged aerosol detector (CAD) using an anion exchange column (Dionex CarboPac PA-1). A new peak was detected (Figure 3.6). This new product was collected and characterized by NMR spectroscopy. The $^{1}H-^{31}P$ heteronuclear multiple quantum correlation (HMQC) spectrum of the substrate, octose 1,8-bisphosphate (**3-2**), displayed two ^{31}P signals. One is at δ 2.53 and is coupled to the two C-8 proton signals at δ 3.86 and δ 3.83, and the other is at δ 0.49 and is coupled to the anomeric proton at δ 5.35 (Figure 3.7). In contrast, the NMR spectrum of the isolated product displayed a single ^{31}P signal at δ 3.17, which is coupled to the anomeric proton at δ 5.34. The loss of the C-8 phosphoryl group from **3-2** during the LmbK-catalyzed reaction is thus evident. To gain more support for the structural assignment of the isolated product, an octose 1-phosphate (**3-3**) standard was also prepared. The LmbK product was found to co-elute with the synthetic octose 1-phosphate (**3-3**) on HPLC, and the spectral characteristics of the enzymatic product are in good agreement with

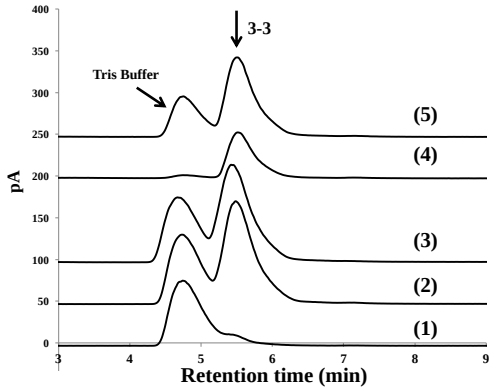


Figure 3.6: HPLC analysis of LmbK activity assays. (1) Control reaction without LmbK. The peak with the retention time of 4.8 min comes from Tris of the protein stock solution. (2) The synthetic bisphosphate **3-2** incubated with LmbK for 2.5 h. (3) The synthetic bisphosphate **3-2** incubated with LmbK for 5 h. (4) The synthetic standard of monophosphate **3-3**. (5) Co-injection of LmbK-reaction (trace 2) and the synthetic standard **3-3**.

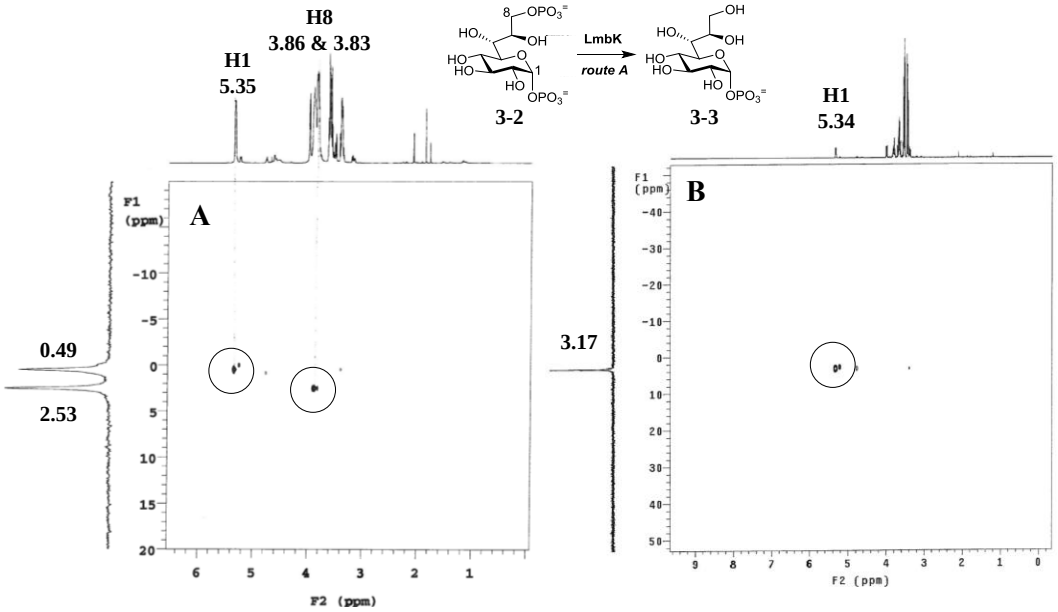


Figure 3.7: NMR characterization of LmbK-catalyzed reaction. The $^{1}H-^{31}P$ HMQC spectra of (A) the synthetic compound **3-2** and (B) the LmbK-catalyzed reaction product **3-3**.

the synthetic standard. These results fully established the identity of the LmbK product as **3-3** starting from **3-2**.

3.3.3 Functional Characterization of LmbO

Nucleotide specificity of LmbO

After the formation of octose 1-phosphate (**3-3**), the next reaction step is likely the nucleotide activation catalyzed by the putative nucleotidylyltransferase, LmbO (Figure 3.1, route A). To test this hypothesis and verify which nucleotide is utilized in the lincomycin biosynthetic pathway, the synthetic bisphosphate **3-2** was incubated with LmbK and the refolded LmbO in the presence of ATP, CTP, GTP, TTP or UTP. The resulting mixtures were analyzed by HPLC (Figure 3.8). Although no new peak was detected in the reaction mixture using ATP, CTP, TTP or UTP, a sample derived from the incubation mixture with GTP showed a new peak that eluted in-between those of GMP and GDP with a retention time of 9.45 min (Figure 3.8, trace 5). The same results were obtained when the synthetic octose 1-phosphate (**3-3**) was incubated with the refolded LmbO (Figure 3.9). This new product peak was collected and subjected to ESI-MS analysis. The recorded molecular mass (calcd for $C_{18}H_{28}N_5O_{18}P_2^-$ $[M-H]^-$, 664.0910; obsd, 664.0922) is consistent with the proposed product, GDP-D-α-D-octose (**3-5**), which was also chemically prepared from **3-3**. As expected, the isolated product from the LmbK/LmbO reaction displayed an identical HPLC retention time as that of the synthetic standard (Figure 3.10, trace 1 and 2).

Substrate specificity of LmbO

These results clearly indicate that the refolded LmbO is catalytically active and suggest a pathway in which C-8 dephosphorylation by LmbK preceeds prior to nucleotidylyltransfer catalyzed by LmbO (Figure 3.1, route A). However, a re-

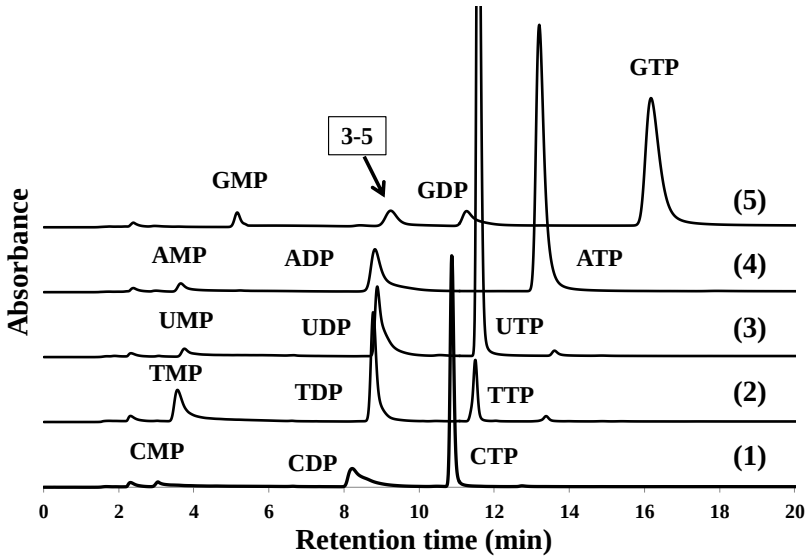


Figure 3.8: HPLC analysis of the acvity assays for LmbK and LmbO. All reactions contained the synthetic bisphosphate **3-2**, LmbK and LmbO. (1) CTP, (2) TTP, (3) UTP, (4) ATP, and (5) GTP.

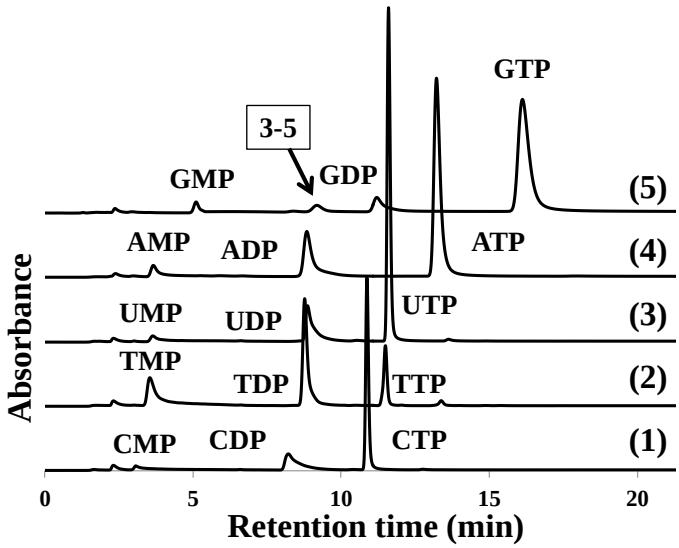


Figure 3.9: HPLC analysis of the activity assays for LmbO. All reactions contained the synthetic monophosphate **3-3** and LmbO. (1) CTP, (2) TTP, (3) UTP, (4) ATP, and (5) GTP.

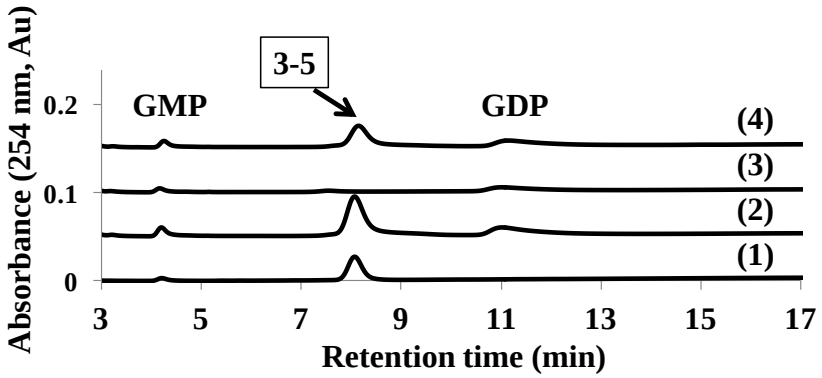


Figure 3.10: Substrate specificity of LmbO. HPLC analysis of the activity assays for LmbO. (1) The synthetic standard of GDP-octose **3-5**. (2) The synthetic bisphosphate **3-2** incubated with GTP, LmbK and LmbO. (3) The synthetic bisphosphate **3-2** incubated with GTP and LmbO. (4) The synthetic monophosphate **3-3** incubated with GTP and LmbO.

versal of the sequence of these two reactions is also conceivable for the biosynthesis of GDP-octose **3-5** (Figure 3.1, route B). To check this possibility, the bisphosphate **3-2** was incubated with the refolded LmbO in the presence of GTP. As shown in Figure 3.10 trace 3, no new peak in the range of GDP-activated sugars is observed. In contrast, formation of GDP-sugar **3-5** (Figure 3.10, trace 4) is clearly visible when the synthetic octose 1-phosphate (**3-3**) was incubated with the refolded LmbO and GTP. Evidently, C-8 dephosphorylation is a prerequisite for the subsequent nucleotidyl activation. These results unequivocally establish that GDP-octose (**3-5**) is indeed an intermediate in MTL biosynthesis and formation of GDP-octose follows route A, not route B.

To investigate whether LmbO tolerates flexibility in its anomeric specificity, the refolded LmbO was incubated separately with GTP and each of the two anomers of glucose 1-phosphate, which were used as substrate analogues. Since formation of GDP-glucose was found only in the sample of α-glucose 1-phosphate (Figure 3.11, trace 1 and 2), LmbO showed stringent α-anomeric stereospecificity for hexose substrates. Interestingly, although LmbO is capable of processing six-carbon

sugar 1-phosphate, LmbK could not hydrolyze α-D-glucose 1,6-bisphosphate (Figure 3.11, trace 3). A similar observation was reported for heptose 1,7-bisphosphate phosphatase GmhB, which could not recognize α-D-glucose 1,6-bisphosphate as a substrate.[131,132] Apparently, the chain length of the sugar phosphate substrates is essential for these phosphatases.

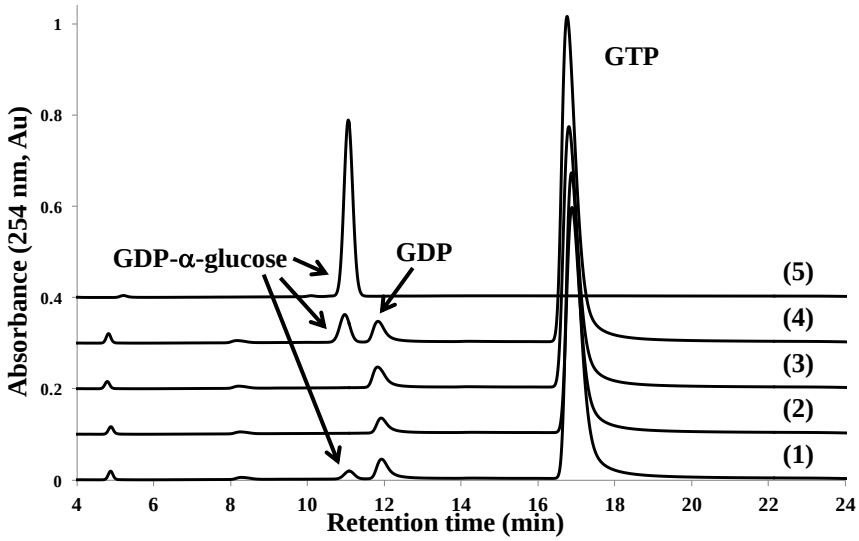


Figure 3.11: Determination of the reaction sequence in NDP-octose biosynthesis. HPLC analysis of the activity assays for LmbK and LmbO. (1) Glucotse α-1-phosphate incubated with GTP and LmbO. (2) Glucotse β-1-phosphate incubated with GTP and LmbO. (3) Glucotse α-1,6-bisphosphate incubated with GTP, LmbK and LmbO. (4) Co-injection of the sample from LmbO reaction with glucose α-1-phosphate (trace 1) and the GDP-α-D-glucose standard. (5) The GDP-α-D-glucose standard.

Moreover, while LmbO was able to attach nucleotide on octose or glucose α-1-phosphate, other hexoses, such as galactose or mannose, were not accepted as substrates (Figure 3.12). This results indicated that the stereochemistry of C2- and C4-hydroxyl groups seems to be crucial for the LmbO-catalyzed nucleotidyl transfer. It is an intriguing observation because GDP-α-mannose and GDP-D-*glycero*-D-*manno*-heptose are the biosynthetic precursors of most naturally occurring GDP-sugars.

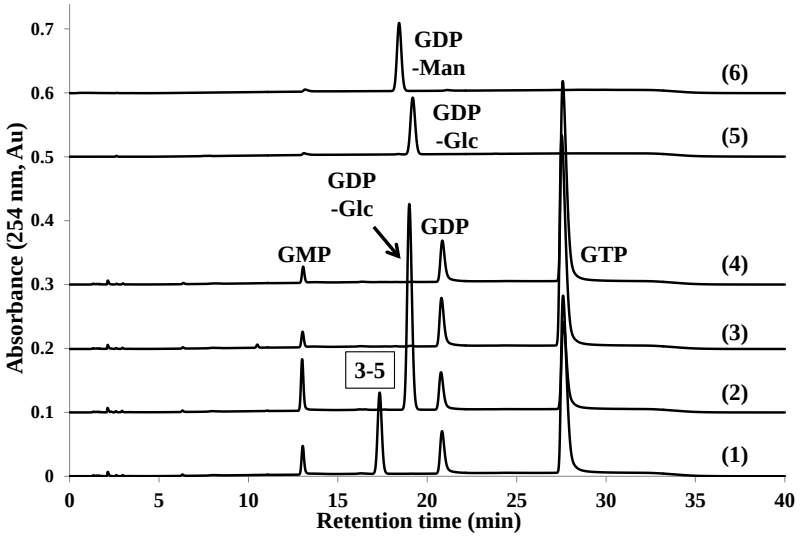


Figure 3.12: Sugar substrate specificity of LmbO. HPLC analysis of the activity assays for LmbO. All reactions contained LmbO and GTP. (1) Octose α-1-phosphate **3-3**, (2) Glucotse α-1-phosphate, (3) Galactose α-1-phosphate, (4) Mannose α-1-phosphate, (5) GDP-α-D-glucose standard, (6) GDP-α-D-mannose standard.

3.4 Conclusion

Taken together, these results provide insight into the lincomycin biosynthetic pathway, part of which is reminiscent of the NDP-heptose pathway.[131,132] The transformation of octose 8-phosphate (**3-1**) to GDP-D-α-D-octose (**3-5**) is shown to involve kinase and phosphatase reactions as intermediary steps. Our data also reveal that the dephosphorylation step, catalyzed by LmbK, is critical for the nucleotide activation reaction. However, direct demonstration of the predicted kinase activity of LmbP was unsuccessful due to the difficulty in refolding insoluble LmbP. Nevertheless, the involvement of a kinase-catalyzed step and the α-stereospecificity of LmbP reaction are supported by the effective reconstitution of the biosynthesis of GDP-D-α-D-octose using the synthetic bisphosphate (**3-2**) as substrate.

The formation of a nucleotide-activated octose intermediate in the lincomycin biosynthetic pathway likely serves two purposes: (1) the phosphonucleotidyl group

might be an important recognition/binding element for the later enzymes in the pathway, and (2) it can function as a good leaving group in a nucleophilic substitution reaction, which may eventually allow the installation of the C-1 thiol group. Furthermore, the identification of the GDP-octose intermediate has settled a long-standing dispute concerning whether the proteins encoded in the *lmb* gene cluster with sequence similarity to some NDP-deoxyhexose modifying enzymes play any roles in lincomycin biosynthesis.[115] It should also be noted that most GDP-activated sugars are used in the biosynthesis of bacterial cell-wall polysaccharides and eukaryotic glycans.[145] Only GDP-mannose has been demonstrated or suggested to be a biosynthetic precursor of the sugar subunit in some secondary metabolites including the polyene macrolide nystatin,[146] amphotericin and candicidin,[147] the aminoglycoside hygromycin,[148] and the antitumor drug bleomycin.[149] Thus, these results identify several key intermediates in the lincomycin pathway and expand our knowledge of the roles of nucleotide-activated sugars in natural product biosynthesis.

Chapter 4

Biosynthetic Studies of Lincomycin A (III): Enzymatic Modifications of the GDP-Octose Intermediate

4.1 Introduction

In the previous chapters, the biosynthetic pathway of lincomycin A was proposed based on the results from early feeding experiments and genetic analysis. The C_8 backbone of thiooctose moiety was demonstrated to be derived from D-sedeheptulose 7-phosphate or D-fructose 6-phosphate and D-ribose 5-phosphate via a transaldol reaction catalyzed by LmbR. Subsequent isomerization catalyzed by LmbN converts the resulting octulose 8-phosphate to octose 8-phosphate. Furthermore, a GDP-octose was identified as a key intermediate in the lincomycin A biosynthetic pathway. The putative phosphatase and guanylyltransferase, LmbK and LmbO, respectively, were found to catalyze the formation of GDP-octose from a bisphosphate intermediate, octose 1,8-bisphosphate.

Compared to TDP- or UDP-activated sugars, the occurrence of GDP-activated sugars in nature is relatively rare. A good example is trehalose 6-phosphate synthase utilizing GDP-D-glucose, instead of other nucleotide diphosphate activated glucoses, has been found in *Streptomyces hydroscopicus*.[150] Most documented enzymatic urilization and/or modifications of GDP-sugars are involved in the biosynthesis of microbial cell-surface polysaccarides. For example, 6-deoxyheptoses have been discovered as part of the lipopolysaccharides or of the capsular polysaccharides on the surface of several pathogens including *Yersinia pseudotuberculosis* and *Campylobacter jejuni*.[145] The biosynthetic precursor of 6-deoxyheptoses has been

found to be GDP-D-*glycero*-α-D-*manno*-heptose (**4-1**), the biosynthesis of which has been described in page 62. It was later shown that two enzymes, DmhA and DmhB, catalyze the sequential C6-dehydration and C4-reduction to produce GDP-6-deoxy-*manno*-heptose (**4-3**) in *Y. pseudotuberculosis* O:2a (Figure 4.1).[151] A similar C6-dehydratase (WcbK) and C4-reductase (WcaG) pair has also been identified in *C. jejuni*.[152] These two enzymes, in addition to Cjj1427 and Cjj1430, are involved in the conversion of GDP-6-deoxy-*manno*-heptose to GDP-6-deoxy-D-*altro*-heptose (**4-1** → **4-5** as dispicted in Figure 4.1).[153]

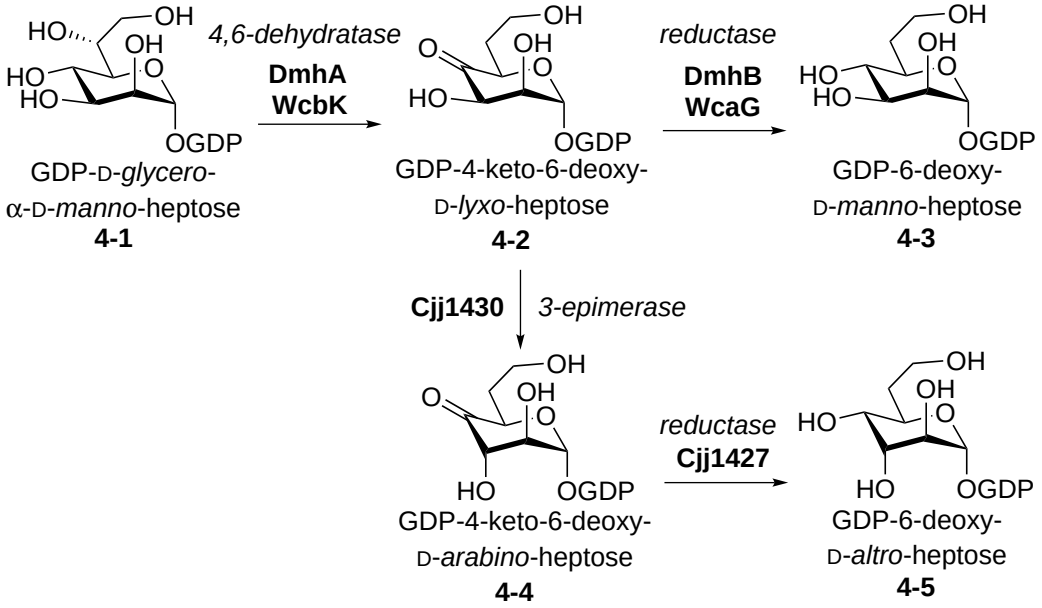


Figure 4.1: Biosynthetic pathway of GDP-6-deoxy-heptose derivatives.

On the other hand, GDP-D-mannose (**4-9**) and its derivatives are hexose donors frequently found in *O*-antigens of Gram-negative bacteria.[145] The biosynthesis of these GDP-sugars starts from a three-step conversion of fructose 6-phosphate (**4-6**) to GDP-D-mannose (**4-9**), which is subsequently converted to GDP-4-keto-6-deoxy-α-D-mannose (**4-10**) via the catalysis of Gmd, a GDP-D-mannose 4,6-dehydratase.[154] This ketosugar **4-10** serves as a central intermediate for the biosynthesis of other GDP-6-deoxyhexoses including GDP-D-rhamnose (**4-11**),[154] GDP-6-

deoxy-D-talose (**4-12**),[155] GDP-perosamine (**4-13**),[156] GDP-L-fucose (**4-14**),[157] and GDP-colitose (**4-15**)[158] (Figure 4.2). GDP-D-mannose has also been demonstrated to be the precursor of vitamin C in plants, although the detailed pathway has yet to be elucidated.[159]

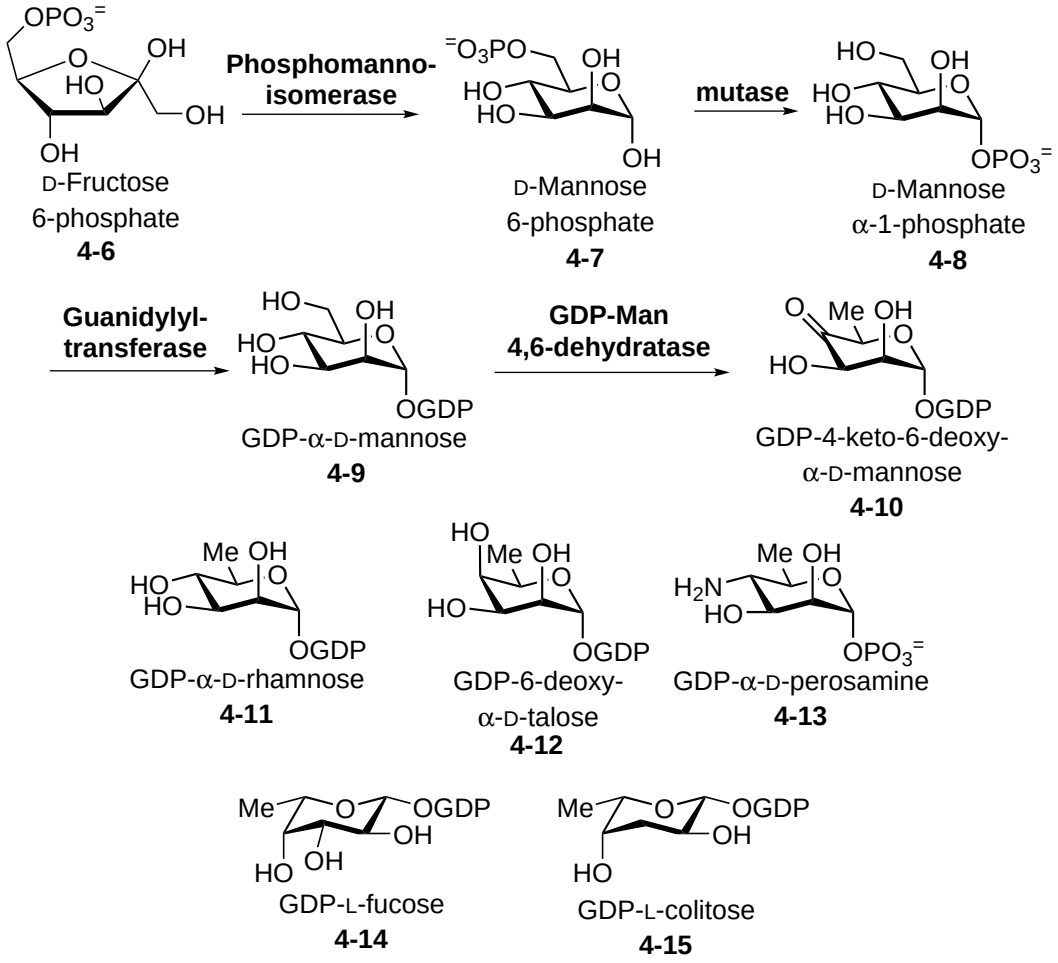


Figure 4.2: Biosynthetic pathway of GDP-α-D-mannose (**4-9**) and other hexose derivatives.

Only a handful of GDP-sugars have been characterized to be the biosynthetic intermediates in secondary metabolites. A well-studied example is the sugar moiety of polyene macrolide antibiotics. Most of the polyene macrolides (including nystatin,[146] amphotericin and candicidin,[147]) contains one single deoxyaminosugar, mycosamine (**4-17**). However, the sugar moiety of perimycin is perosamine

(**4-13**), a constitutional isomer of mycosamine (Figure 4.3). Both deoxyaminosugars have been proposed to be biosynthesized from GDP-α-D-mannose (**4-9**), which could be converted to GDP-4-keto-6-deoxy-α-D-mannose (**4-10**) by GDP-mannose 4,6-dehydratase in a similar manner as that observed in the biosynthesis of bacterial *O*-antigens.[146] The C4-ketosugar **4-10** then undergoes transamination to produce GDP-perosamine (**4-13**).[160] For the biosynthesis of GDP-mycosamine (**4-17**), the C4-ketosugar **4-10** is isomerized to GDP-3-keto-6-deoxymannose (**4-16**), which serves as the putative substrate for the C3-aminotransferase.[161] However, no homologous genes encoding GDP-4-keto-6-deoxymannose 3,4-isomerase have been found in the polyene macrolide biosynthetic gene clusters. It is possible that the isomerization is a non-enzymatic reaction or catalyzed by other cytosolic proteins.

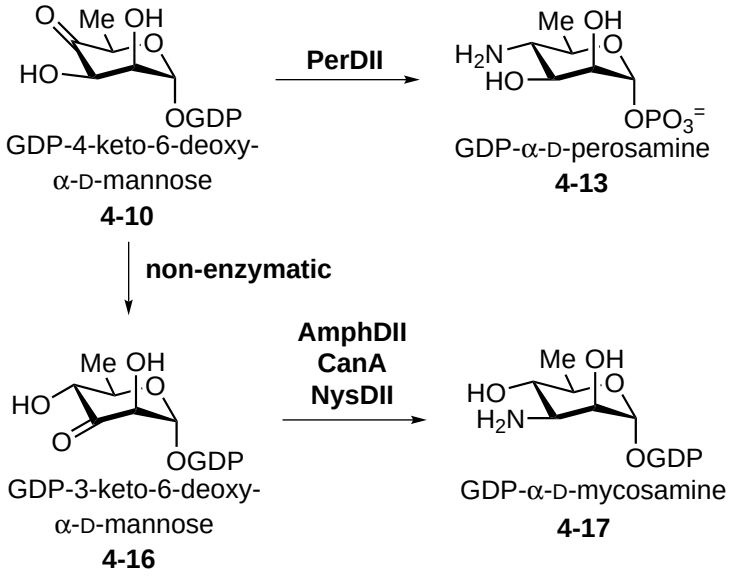


Figure 4.3: Biosynthetic pathway of the sugar moieties in polyene macrolide antibioties.

In the biosynthetic pathway of lincomycin A, the attachment of a guanosine-diphosphate group at C-1 position may allow the octose intermediate to be recognized and further utilized by the GDP-sugar modifying enzymes. Comparing the chemical structures of lincomycin A and the GDP-octose intermediate, three possible modifications could be envisioned: C-4 epimerization, C-6 transamination and

C-8 dehydration (Figure 4.4).

Based on the bioinformatic analysis of the lincomycin A biosynthetic gene cluster, LmbM (an epimerase homolog) was proposed to catalyze the C-4 epimerization. Following the reaction mechanism similar to that of NDP-hexose 4,6-dehydratase, the C-8 dehydration is likely accompanied with the C-6 dehydrogenation catalyzed by the nicotinamide adenine dinucleotide (NAD^+) or nicotinamide adenine dinucleotide phosphate ($NADP^+$)-dependent enzymes.[11] Two protein candidates encoded in the *lmb* cluster are LmbL (a putative NDP-sugar dehydrogenase) and LmbZ (a putative oxidoreductase). Although the sequence of these enzymatic modifications is uncertain, the C-8 dehydration is predicted to occur prior to the C-6 transamination because the presence of a keto-group is a prerequisite for the PLP-dependent transamination.[162] Search of the conserved domains within *lmbS* gene product led to the annotation of its encoded protein as a putative NDP-ketosugar aminotransferase, which belongs to the DegT/DnrJ/EryC1/StrS aminotransferase family. Three biosynthetic routes of GDP-octose modifications with the presumed stereochemical changes are summarized in Figure 4.4.

4.2 Experimental Procedures

4.2.1 General

Materials

All chemicals and reagents were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA) and were used without further purification unless otherwise specified. Oligonucleotide primers were prepared by Integrated DNA Technologies (Coralville, IA). Kits for DNA gel extraction and spin minipreps were purchased from Qiagen (Valencia, CA). PureLink® Genomic DNA Mini Kit was obtained from from Invitrogen (Carlsbad, CA). KOD NDA polymerase was purchased from Novagen (Madison, WI). QickChange site-

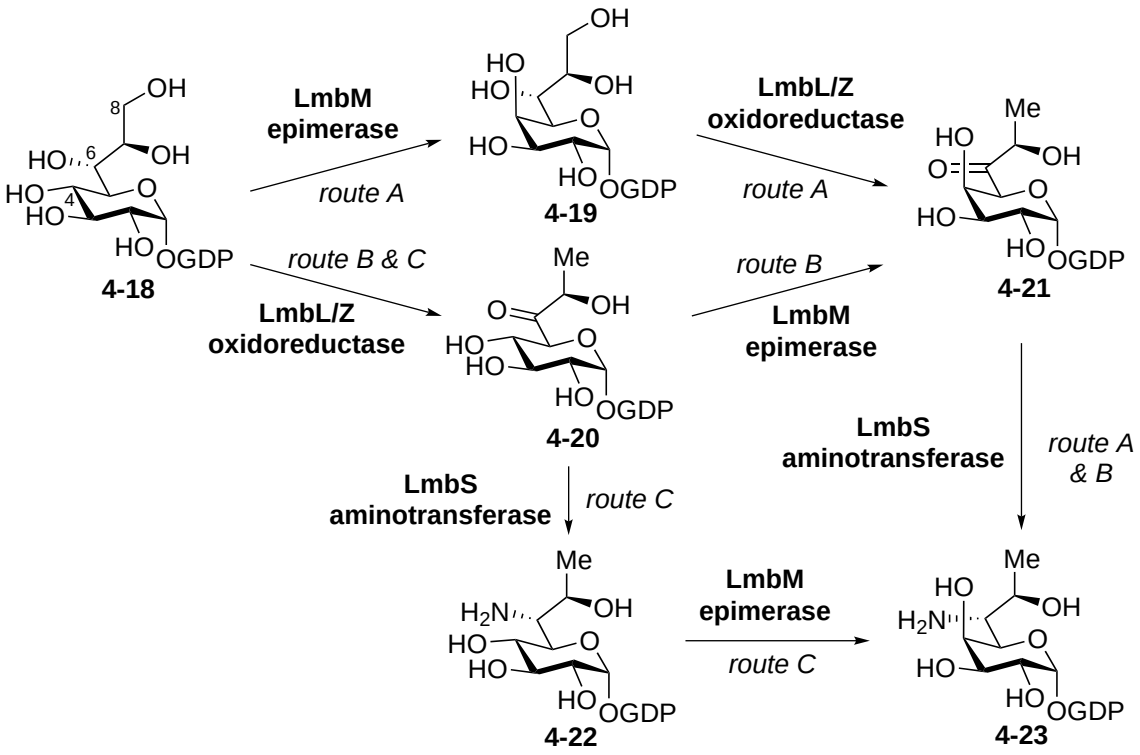


Figure 4.4: Proposed biosynthetic pathway of the GDP-octose modifications.

directed mutagenesis kit was obtained from Stratagene, which was acquired by Agilent (Santa Clara, CA). Enzymes and molecular weight standards used for the cloning experiments were obtained from New England Biolabs (Ipswich, MA). Reagents for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) were purchased from Bio-Rad (Hercules, CA), with the exception of the protein molecular weight markers, which were obtained from Invitrogen. Growth medium components were acquired from Becton Dickinson (Sparks, MD). Sterile syringe filters are products of Fisher Scientific. Amicon YM-10 ultrafiltration membranes were bought from Millipore (Billerica, MA). The analytical and semi-preparative CarboPac PA1 high-performance liquid chromatography (HPLC) columns were obtained from Dionex (Sunnyvale, CA). Analytical C-18 HPLC columns are products of Varian (Palo Alto, CA). Simi-preparative C-18 HPLC columns were purchased from Fisher Scientific (Pittsburgh, PA).

Bacterial Strains and Plasmids

Streptomyces lincolnensis NRRL ISP-5355 (identical to ATCC 25466) and *Streptomyces caelestis* NRRL 2418 were obtained from the Agricultural Research Service (ARS) Culture Collection of the National Center for Agricultural Utilization Research (Peoria IL). *Escherichia coli* DH5α, acquired from Bethesda Research Laboratories (Gaithersburg, MD), was used for routine cloning experiments. The protein overexpression host *E. coli* BL21 star (DE3) was obtained from Invitrogen. Vector pET28b(+) for protein overexpression was purchased from Novagen (Madison, WI).

Instrumentation

Standard genetic manipulations of *E. coli* were performed as described by Sambrook and Russell.[142] DNA sequencing was performed at the core facility of the Institute of Cellular and Molecular Biology, The University of Texas at Austin. Vector NTI Advance 10.1.1 from Invitrogen was used for sequence alignments. DNA concentrations were measured using a NanoDrop ND-1000 UVvis instrument from Thermo Fisher Scientific. HPLC was performed on a Beckman Coulter System Gold HPLC equipped with a UV detector or a Corona CAD (charged aerosol detector, ESA Biosciences, Chelmsford, MA). NMR spectra were acquired using a Varian Unity 400 MHz, 500 MHz or 600 MHz spectrometer, and chemical shifts (in ppm) are reported relative to that of the solvent peak (δ_H = 7.24 and δ_C = 77.0 for deuterated chloroform, δ_H = 4.67 for deuterium oxide). Mass spectra were recorded at the Mass Spectrometry core facility in the Department of Chemistry and the Proteomics Facility in College of Pharmacy at the University of Texas, Austin. Optical rotations were measured using an ATAGO AP-300 automatic polarimeter at a path length of 1 dm.

4.2.2 Cloning of *lmbL*, *lmbM*, *lmbS*, *lmbZ*, *ccbS* and *ccbZ*

Table 4.1: Primers used for constructing plasmids containing *lmbL*, *lmbM*, *lmbS*, *lmbZ*, *ccbS* and *ccbZ* genes.

Plasmid	Primer pair
pET24b(+)-*lmbL*-CHis_6	CIL015 / CIL016
pET28b(+)-*lmbM*-CHis_6	CIL017 / CIL018
pET24b(+)-*lmbS*-CHis_6	CIL019 / CIL020
pET28b(+)-*lmbS*-NHis_6	CIL019 / CIL021
pET24b(+)-*lmbZ*-CHis_6	CIL022 / CIL023
pET28b(+)-*lmbZ*-NHis_6	CIL022 / CIL024
pET24b(+)-*ccbS*-CHis_6	CIL025 / CIL026
pET28b(+)-*ccbS*-NHis_6	CIL025 / CIL027
pET-MalE-*ccbS*	CIL025 / CIL027
pET24b(+)-*ccbZ*-NHis_6	CIL028 / CIL029

Primer number	Primer name	Sequence[a]
CIL015	F-*lmbL*-24/28-NdeI	5′-AGTTAAG**CATATG**ACCGACGCGACGCACC-3′
CIL016	R-*lmbL*-24-HindIII	5′-ATAT**AAGCTT**CGGGGCGGCATGCG-3′
CIL017	F-*lmbM*-24/28-NdeI	5′-AGTTAAG**CATATG**AGCGGGCGGTACTGC-3′
CIL018	R-*lmbM*-24-HindIII	5′-ATAT**AAGCTT**CGGCCGCTGCCACC-3′
CIL019	F-*lmbS*-24/28-NdeI	5′-GGCTAAG**CATATG**AGCGACTACATCCCCTTCG-3′
CIL020	R-*lmbS*-24-HindIII	5′-ATAA**AAGCTT**CCGTCCCGCCTCGTCG-3′
CIL021	R-*lmbS*-28-HindIII	5′-ATAA**AAGCTT***TCA*CCGTCCCGCCTCG-3′
CIL022	F-*lmbZ*-24/28-NdeI	5′-ATCACAA**CATATG**ACCCACAGGTGCGGC-3′
CIL023	R-*lmbZ*-24-HindIII	5′-TGAT**AAGCTT**CGGTTTCTCCCAGGTGAGG-3′
CIL024	R-*lmbZ*-28-HindIII	5′-ACAT**AAGCT*T***CACGGTTTCTCCCAGGTGAGG-3′
CIL025	F-*ccbS*-24/28-NdeI	5′-TCTATAA**CATATG**AGCGACTACGTCCCCTTTGC-3′
CIL026	R-*ccbS*-24-HindIII	5′-ATAT**AAGCTT**GCTCCTCGTGCCCTCG-3′
CIL027	R-*ccbS*-28-HindIII	5′-TTAC**AAGCT*T***CAGCTCCTCGTGCCCTCG-3′
CIL028	F-*ccbZ*-24/28-NdeI	5′-TCTATAA**CATATG**ACATCCCCCACCCGC-3′
CIL029	R-*ccbZ*-28-HindIII	5′-TTAC**AAGCT*T***CACGTGCTGGTGCGG-3′

[a]The engineered restriction sites are shown in bold; the start codon is shown in bold and underlined; the stop codon is shown in italic.

Plasmids containing *lmbL*, *lmbM*, *lmbS*, *lmbZ*, *ccbS* and *ccbZ* were constructed by amplifying the corresponding genes from the genomic DNA isolated from *S. lincolnensis* and *S. caelestic* using designed primer pairs (Table 3.1). The PCR products were then digested with NdeI and HindIII and ligated into pET24b(+) and pET28b(+), which had been digested with the same restriction enzymes. The *ccbS*

gene was also sub-cloned into the pET-MalE vector (generated in house) to obtain the nearly native protein form (i.e., possessing two additional amino acids, Gly-His, at the *N*-terminus) after TEV cleavage. The resulting plasmids were sequenced using T7 or T7 terminal universal primer and used to transform the *E. coli* BL21 star (DE3) for protein overexpression.

4.2.3 Overexpression and Purification of His-tagged Proteins in *E. coli*

Overexpression and purification procedures for *C*-His_6-LmbL, *C*-His_6-LmbM and *N*-His_6-CcbZ were identical to those used to purify *N*-His_6-LmbK as described in Sections 3.2.4 and 3.2.5. The yields of the *N*-His_6 tagged proteins from 1 L culture were *C*-His_6-LmbL (∼40 mg), *C*-His_6-LmbM (∼18 mg) and *N*-His_6-CcbZ (∼16 mg). The molecular mass and purity of the purified proteins were determined by SDS-PAGE analysis

4.2.4 Overexpression and Purification of CcbS from Maltose Binding Protein Fusion System

The CcbS protein without His_6-tag was obtained by *in vitro* TEV enzyme cleavage of the maltose binding protein (MBP)-CcbS fusion protein. A standard procedure for protein overexpression as described in Sections 3.2.4 was applied. However, for the induction of protein expression, the amount of IPTG was adjusted to 50 μM. The thawed cells (∼20 g) were resuspended in the lysis buffer (100 mL) containing 50 mM Tris · HCl (pH 8), 10% (v/v) glycerol and 10 mM imidazole. The mixture was added PLP to 1 mM as the final concentration. After incubation with lysozyme (100 mg) for 1 h, the cells were disrupted by sonication using 10 × 10-s pulses with a 20-s cooling pause between each pulse. The resulting lysate was centrifuged at 15,000×g for 30 min, and the supernatant was incubated with 10 mL of the lysis buffer pre-equilibrated Ni-NTA resin for 16 h. The mixture

was then loaded onto a column, drained, and washed with wash buffer containing 50 mM Tris · HCl (pH 8), 300 mM NaCl, 20 mM imidazole and 10% glycerol until the eluent became colorless. Bound protein was eluted using 50 mM Tris · HCl buffer (pH 8) containing 250 mM imidazole and 10% glycerol. The column was then loaded with 1 mL more of the Ni-NTA resins and re-equilibrated with the lysis buffer to be ready for the subsequent purification. The eluted proteins were dialyzed against 3 × 1-L of the dialysis buffer containing 50 mM Tris · HCl (pH 8), 300 mM NaCl and 15% glycerol. After 2 hours of the first time dialysis, 5% (v/v) His_6-tagged TEV protease was added to the solution containing MBP fusion proteins to cleave the His_10-MBP. The digestion was performed for 24 h during dialysis. The protein mixture was then passed through the above-mentioned Ni-NTA resins dropwise. The flow through was passed through the column one more time to ensure the complete removal of His-tagged MBP and TEV. The Ni-NTA resins were then washed with 2-column volume of wash buffer. The collected protein fractions were combined and concentrated with an Amicon ultra-15 centrifugal filter unit with 10 kDa cut-off. The yield of CcbS was approximately 5 mg from 1 L culture. The molecular mass and purity of CcbS were determined by SDS-PAGE analysis.

4.2.5 Identification of Cofactors Bound to LmbL, LmbM and CcbZ

To denature the proteins, the purified *C*-His_6-LmbL protein (4.26) and *C*-His_6-LmbM (2.39 nmole) in 350 μL of 50 mM Tris buffer (pH 8) were heated to 90 °C in a water bath for 5 min. The purified *N*-His_6-CcbZ (1.64 nmole) in 50 μL of 50 mM Tris buffer (pH 8) was treated with 300 μL of methanol. Each of the denatured protein solution was then centrifuged at 16,000×*g* for 20 min (4 °C) to remove the precipitates, and the supernatant was lyophilized to near dryness and then re-dissolved in 50 μL of water. All samples were analyzed by HPLC using a

Dionex CarboPac PA1 analytical column (4 × 250 mm) with water as mobile phase A was water and 1 M NH_4OAc as mobile phase B. The separation was conducted at a flow rate of 1 mL/min with a gradient program: 0-2 min 10% B, 2-10 min 10-50% B, 10-25 min 50-90% B, 25-30 min 90% B, 30-32 min 90-10% B, 32-40 min 10%B. The detector wavelength was set at 259 nm.

4.2.6 Transamination Half-Reactions of CcbS

The half-reaction mixture contained CcbS protein (46.8 μM), L-amino acid (1 mM) in 50 mM Tris buffer (pH 8). The mixture was incubated at ambient temperature for 3 h before spectro-photoscanning by UV-Vis spectrophotometer. Amino acids include L-alanine, L-aspartate, L-glutamine and L-glutamate.

4.2.7 LmbL, LmbM, CcbS and CcbZ Reaction Assays

Functional Characterization of LmbL, LmbM and CcbZ

The purified GDP-octose **4-18** (1 mM) was incubated with *C*-His_6-LmbL protein (4.26 μM final concentration), NAD^+ (1.25 mM) and $MgCl_2$ (10 mM) in 50 mM Tris buffer (pH 8.0) at ambient temperature for 6 h. Reactions containing *C*-His_6-LmbM (4.78 μM) and/or *N*-His_6-CcbZ (3.27 μM) were were similarly prepared. At the end of incubation, each reaction mixture was filtered through a YM-10 membrane to remove the protein. HPLC analysis of the filtrate was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm) with water as mobile phase A and 1 M NH_4OAc as mobile phase B. The separation was conducted at a flow rate of 1 mL/min with a gradient program: 0-2 min 60% B, 2-14 min 60-90% B, 14-24 min 90% B, 24-26 min 90-60% B, 26-34 min 60% B. The detector wavelength was set at 254 nm.

NMR characterization of LmbM product

The reaction solution (0.3 mL × 5) containing compound **4-18** (5 mM), NAD^+ (0.125 mM) and *C*-His_6-LmbM (4.78 M) in water was incubated at ambient temperature for 16 h and then filtered through a YM-10 membrane using an Amicon ultrafiltration unit to remove proteins. The filtrate was then purified using HPLC using a Dionex CarboPac PA1 semi-prep column (9 × 250 mm) with water as mobile phase A and 1 M NH_4OAc as mobile phase B. The separation was conducted at a flow rate of 3 mL/min with a gradient program: 0-2 min 60% B, 2-14 min 60-90% B, 14-24 min 90% B, 24-26 min 90-60% B, 26-34 min 60% B. The detector wavelength was set at 254 nm. The retention time of the LmbM-product, GDP-D-*erythro*-α-D-*galacto*-octose (**4-19**), was ∼19 min. ^{1}H NMR (D_2O, 600 MHz) δ (ppm) 7.99 (1 H, s, guanine), 5.82 (1 H, d, $J_{1,2}$ = 5.4 Hz, ribose-H1), 5.46 (1 H, dd, $J_{1,2}$ = 3.6 Hz, $J_{1,P}$ = 7.2 Hz, octose-H1), 4.60 (1 H, dd, $J_{1,2}$ = $J_{2,3}$ = 5.4 Hz ribose-H2), 4.37 (1 H, dd, $J_{2,3}$ = 5.4 Hz, $J_{3,4}$ = 4.2 Hz, ribose-H3), 4.23 (1 H, brs, ribose-H4), 4.23–4.06 (2 H, m, ribose-H5), 3.89 (1 H, dd, $J_{5,6}$ = 1.2 Hz, $J_{6,7}$ = 6.6 Hz, octose-H6), 3.84 (1 H, ddd, $J_{6,7}$ = 6.6 Hz, $J_{7,8a}$ = 3.6 Hz, $J_{7,8b}$ = 6.6 Hz, octose-H7), 3.82 (1 H, dd, $J_{4,5}$ = 9.6 Hz, $J_{5,6}$ = 1.2 Hz, octose-H5), 3.73 (1 H, dd, $J_{7,8a}$ = 3.6 Hz, $J_{8a,8b}$ = 12.0 Hz, octose-H8$_a$), 3.71 (1 H, brd, $J_{2,3}$ = 10.2 Hz, octose-H3), 3.60 (1 H, dd, $J_{7,8b}$ = 6.6 Hz, $J_{8a,8b}$ = 12.0 Hz, octose-H8$_b$), 3.58 (1 H, brd, $J_{4,5}$ = 9.6 Hz, octose-H4), 3.52 (1 H, m, octose-H2). ^{31}P NMR (D_2O, 243 MHz) δ (ppm) −11.34 (19.0 Hz), −3.12 (19.0 Hz). ESI-HRMS Calcd. for $C_{18}H_{28}N_5O_{18}P_2[M-Na]^-$ 664.0910; Found: 664.6908.

LmbL and CcbZ reaction using GDO-octose 4-19 as substrate

The purified GDP-octose **4-19** (0.5 mM) was incubated with *C*-His_6-LmbL protein (9 μM), *N*-His_6-CcbZ protein (9 μM) and NAD^+ (0.2 mM) in 30 mM sodium phosphate buffer (pH 7.0) at 30 °C. Control experiments containing only LmbL or CcbZ were similarly prepared. Aliquots were withdrawn at different time inter-

vals and then filtered through a YM-10 membrane to remove the protein. HPLC analysis of the filtrate was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm) with water as mobile phase A and 1 M NH_4OAc as mobile phase B. The separation was conducted at a flow rate of 1 mL/min with a gradient program: 0-2 min 10% B, 2-10 min 10-50% B, 10-25 min 50-90% B, 25-30 min 90% B, 30-32 min 90-10% B, 32-40 min 10%B. The detector wavelength was set at 254 nm. Enzymatic assays described below were analyzed by HPLC using the same separation program unless further specified.

LmbL, LmbM and CcbZ reaction using GDO-octoses 4-18 and 4-19 as substrates

The purified GDP-octose **4-18** or **4-19** (0.5 mM) was incubated with *C*-His_6-LmbL protein (9 μM), *C*-His_6-LmbM protein (9 μM), *N*-His_6-CcbZ protein (9 μM) and NAD^+ (0.2 mM) together in 30 mM sodium phosphate buffer (pH 7.0) at 30 °C. The filtrates after the removal of proteins were analyzed by HPLC.

CcbS reaction using ketosugars 4-24 and 4-25 as substrates

The ketosugars **4-24** and **4-25** were generated using the condition as described in previous section. The reaction mixtures were filtered through a YM-10 membrane to remove proteins. The filtrates were then incubated with CcbS (20 μM), L-glutamate (3 mM) and PLP (0.5 mM) at 30 °C for 2 h. After the removal of proteins, the filtrates were analyzed by HPLC.

Reverse transamination of CcbS using the synthetic aminooctose 4-23 as substrate

The synthetic GDP-aminooctose **4-23** (0.2 mM) was incubated with CcbS (12 μM), α-ketoglutarate (6 mM), PLP (0.12 mM) in 100 mM sodium phosphate buffer (pH 7.0) at 30 °C. Aliquots were withdrawn at different time intervals and

then filtered through a YM-10 membrane to remove proteins. HPLC analysis of the filtrate was performed using the similar method as described above.

4.3 Results and Discussion

4.3.1 Overexpression and Purification of LmbL, LmbM, LmbS/ CcbS and LmbZ/CcbZ in *E. coli*

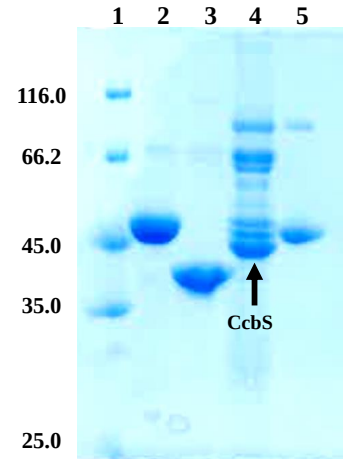


Figure 4.5: SDS-PAGE gel of the purified proteins: *C*-His$_6$-LmbL (47.4 kDa, lane 2), *C*-His$_6$-LmbM (37.0 kDa, lane 3), Gly-His-CcbS (41.9 kDa, lane 4) and *N*-His$_6$-CcbZ (48.0 kDa, lane 5). The molecular markers are shown in lane 1. Co-elution of minor impurities were possibly resulted from the non-specific interactions with Ni-NTA resins.

To explore further modifications of GDP-octose in lincomycin A biosynthesis, the encoding genes, *lmbL*, *lmbM*, *lmbS* and *lmbZ*, were amplified and cloned into pET24b(+) or pET28b(+) vectors. The recombinant *C*-His$_6$ tagged LmbL and LmbM could be purified as soluble proteins (Figure 4.5, lane 2 and lane 3), but LmbS and LmbZ were expressed only as inclusion bodies. Since CcbS and CcbZ encoded in the *ccb* gene cluster are expected to serve similar roles as LmbS and LmbZ, respectively, in both lincomycin A and celesticetin biosynthetic pathways, *ccbS* and *ccbZ* genes were also amplified and cloned into pET24b(+), pET28b(+) or pET-MalE vectors. The *N*-His$_6$-CcbZ was obtained in soluble form (Figure 4.5, lane 5).

While both *C*- and *N*-His_6-CcbS could be overexpressed in soluble forms, the protein yields poor due to slow growth rate and low cell density after IPTG induction (∼1 g of cells per liter). Moreover, the purified His_6-tagged CcbS proteins were colorless instead of lightly yellow. Considering that CcbS was predicted to be a PLP-dependent aminotransferase, the colorless of the as-purified His_6-tagged CcbS indicated the absence of PLP cofactor, which might be resulted from improper protein folding. Thus, the MBP-CcbS construct was prepared and expressed in *E. coli* BL21 star DE3. Growing of the recombinant strain gave improved cell density (∼8 g of cells per liter). The soluble MBP-CcbS fusion proteins were purified and cleaved by TEV enzyme using standard procedure. Excess PLP was added to the lysis buffer to aid the folding process of CcbS. Although multiple bands were detected in the fractions containing CcbS (Figure 4.5, lane 4), the as-purified enzymes showed distinct yellow color indicating the presence of PLP.

4.3.2 Identification of Cofactor Bound to LmbL, LmbM and CcbZ

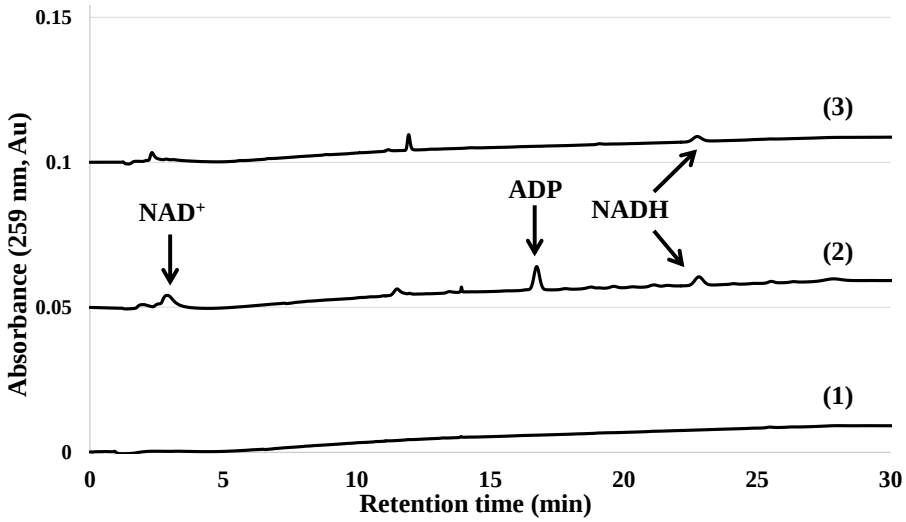


Figure 4.6: HPLC analysis of cofactor bound to (1) LmbL, (2) LmbM, and (3) CcbZ.

Since LmbL, LmbM and CcbZ all belong to the short-chain dehydrogenase/reductase (SDR) family. Members of this family share a conserved binding

fold for NAD(P)$^+$ and are involved in a variety of reactions, including ketoreduction, epimerization, oxidation/dehydration, α-epimerization/ketoreduction, and oxidation/decarboxylation. In order to determine whether the as-purified *C*-His$_6$-LmbL, *C*-His$_6$-LmbM and *N*-His$_6$-CcbZ contain the predicted nicotinamide cofactor, these proteins were precipitated by heating or treating with methanol. The supernatant from each proteins was then lyophilized and analyzed by HPLC. As shown in Figure 4.6, both NAD$^+$ and NADH were found to be bound to LmbM (trace 2). While NADH was observed in the supernatant of denatured CcbZ (trace 3), no nicotinamide cofactor was detected in the extract of LmbL (trace 1).

4.3.3 Transamination Activity of CcbS with Different Amino Donors

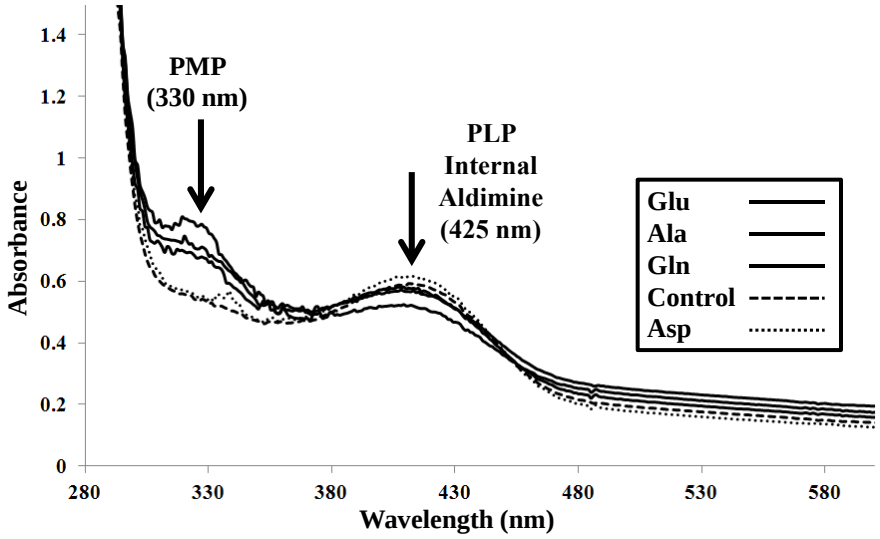


Figure 4.7: Half-reaction of the PLP-dependent aminotransferase CcbS. Control reaction omitting the addition of amino acid was run in parallel.

The PLP-dependent transamination can be divided into two half-reactions. The first step is the transfer of the amino group from an amino acid donor to PLP to generate pyridoxamine 5′-phosphate (PMP) in the active site. The PMP cofactor is then utilized to transfer the amino group to a keto-group-containing acceptor. The purified CcbS was tested for its capability to carry out the first half-reaction of

transamination using different amino donors, L-alanine, L-aspartate, L-glutamate and L-glutamine (Figure 4.7). The putative aminotransferase CcbS showed the highest activity when L-glutamate was used as the amino donor. Lower activity was also observed when CcbS was incubated with L-alanine or L-glutamine. Hence, the half-reaction activity of CcbS purified from the MBP-fusion was confirmed, and L-glutamate is likely the amino acid donor for CcbS-catalyzed transamination reaction.

4.3.4 Functional Characterization of LmbL, LmbM, CcbS and CcbZ

LmbM as the C-4 epimerase

To determine the sequence of the modification reactions, the GDP-octose **4-18** was incubated with LmbL, LmbM or CcbZ, separately or different combination of these three enzymes. HPLC analysis showed that there was no obvious consumption of **4-18** in the reaction with LmbL or CcbZ (Figure 4.8, trace 2 and 4). However, in the LmbM reaction, the substrate **4-18** was fully consumed and a new peak with a shorter retention time was observed (trace 3). This new peak was collected from multiple injections and subjected to ESI-MS analysis. The recorded molecular mass (calcd for $C_{18}H_{28}N_5O_{18}P_2^-$ $[M-H]^-$, 664.0910; obsd, 664.0908) is identical to that of substrate **4-18** or its stereoisomers.

The LmbM-product was further characterized by NMR spectroscopy. Although no removal of hydroxyl group or formation of methyl group was observed, the coupling constant between H3–H4 of LmbM-product becomes smaller as a stereochemical change of the octose ring structure. The observed small $J_{3,4}$ and a large $J_{4,5}$ are similar to those of the chemically synthesized *galacto*-octose standard (Section 2.2.5, page 42). While a smaller coupling constant between H4–H5 is expected, the extension of the carbon chain may potentially twist the conformation of the pyranose ring and cause the increase of the coupling constant. Thus,

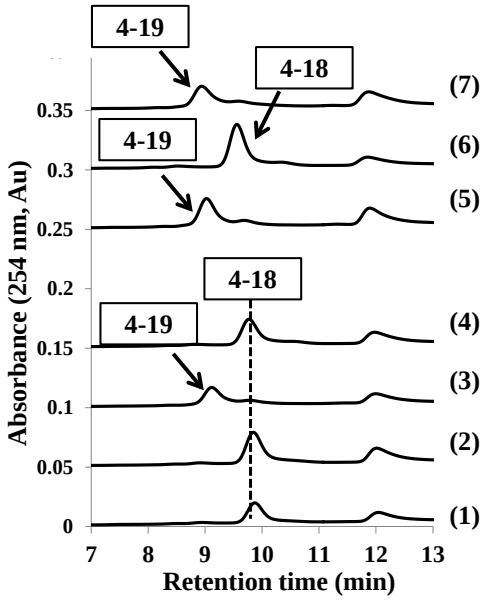


Figure 4.8: HPLC analysis of LmbM, LmbL and CcbZ reactions using **4-18** as the substrate. (1) Control reaction containing substrate **4-18** with no enzyme added. Substrate **4-18** incubated with (2) LmbL, (3) LmbM, (4) CcbZ, (5) LmbL and LmbM, (6) LmbL and CcbZ, (7) LmbM and CcbZ.

the product of LmbM reaction can be assigned as GDP-D-*glycero*-D-*galacto*-octose (**4-19**), the C-4 epimer of **4-18**. These observations indicate that the first step of the GDP-octose modification in lincomycin A biosynthesis is the LmbM-catalyzed C-4 epimerization (Figure 4.4 route A).

Octose dehydration requires both LmbL and CcbZ

The purified LmbM-product (**4-19**) was then incubated with LmbL and/or CcbZ in order to identify the putative dehydratase. The results are shown in Figure 4.9. The sugar substrate **4-19** remained practically unchanged upon incubation with LmbL or CcbZ (trace 1 and 3). In contrast, significant consumption of the substrate **4-19** was observed when both LmbL and CcbZ were added to the enzymatic reaction (trace 2, 4, 5 and 6). The formation of a new peak with a retention time around 20 min (**4-24**) was independent of the addition order of enzymes. This product was later collected and subjected to ESI-MS analysis. The recorded molec-

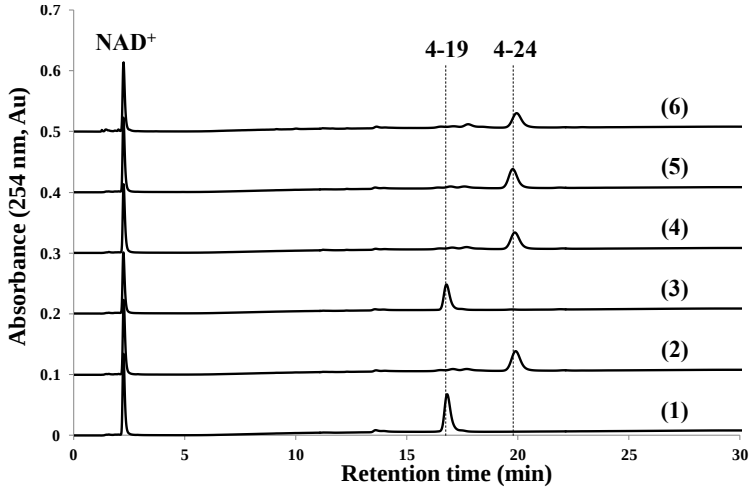


Figure 4.9: HPLC analysis of LmbL and CcbZ reactions using **4-19** as the substrate. All reaction mixtures contained NAD^+ and **4-19** and were incubated at 30 °C. (1) Incubated with only LmbL for 1 h. (2) Incubated with LmbL for 1 h, and then added CcbZ for an additional 2-h incubation. (3) Incubated with only CcbZ for 1 h. (4) Incubated with CcbZ for 1 h, and then added LmbL for an additional 2-h incubation. (5) Incubated with LmbL and CcbZ for 1 h. (6) Incubated with LmbL and CcbZ for 3 h.

ular mass (calcd for $C_{18}H_{27}N_5O_{17}P_2^-$ $[M-H]^-$, 646.0804; obsd, 646.0809) matched the expected number of a dehydrated GDP-ketosugar (see Figure 4.14 for possible structure of **4-24**).

However, when the sugar **4-18**, the substrate of LmbM, was incubated with the three oxidoreductases, LmbM, LmbL and CcbZ, formation of the ketosugar **4-24** was not observed. Instead, the substrate **4-18** was converted to a different product (**4-25**) with a retention time of ∼18 min under the reaction condition (Figure 4.10, trace 3). Based on ESI-MS analysis, the molecular mass of this new product appeared to be 646.0800, which is identical to that of **4-24** and is also consistent with that of a dehydrated product (calcd. for $C_{18}H_{27}N_5O_{17}P_2^-$ $[M-H]^-$, 664.0800) (see Figure 4.14 for possible structure of **4-25**).

A separate experiment in which the epimerization product **4-19** was treated with all three enzymes also yielded the ketosugar **4-25** instead of **4-24** (Figure 4.11, trace 3 and 4). Another peak with the retention time close to **4-25** was observed

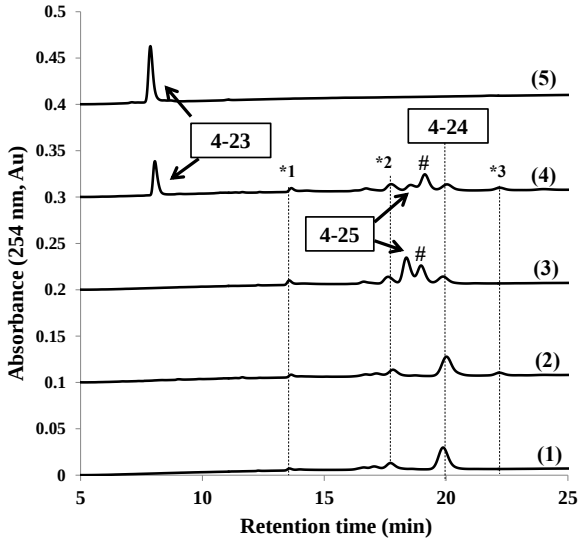


Figure 4.10: HPLC analysis of dehydration reaction catalyzed by LmbL and CcbZ using **4-18** or **4-19** as substrates. All reaction mixtures contained NAD^+. (1) GDP-sugar **4-19** incubated with LmbL and CcbZ. (2) Reaction mixture of (1) was filtered to remove proteins, and then incubated with CcbS, PLP and L-glutamate. (3) GDP-sugar **4-18** incubated with LmbL, LmbM and CcbZ. (4) Reaction mixture of (3) was filtered to remove proteins, and then incubated with CcbS, PLP and L-glutamate. ($\star$ 1: GMP, $\star$ 2: residual **4-18**, $\star$ 3: NADH, #: a possible side product (**4-30**) due to over-reduction.) (5) Synthetic standard GDP-aminooctose **4-23**.

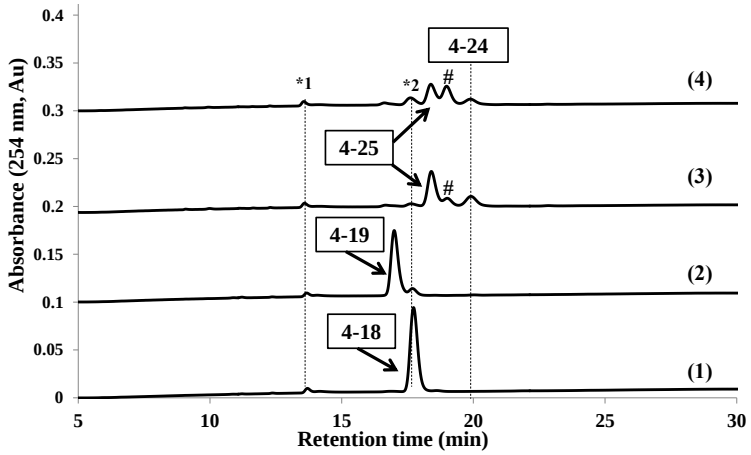


Figure 4.11: HPLC analysis of dehydration reaction catalyzed by LmbL and CcbZ using **4-19** as substrates. (1) GDP-octose **4-18**. (2) GDP-octose **4-19**. (3) GDP-sugar **4-19** incubated with LmbL, LmbM and CcbZ. (4) Reaction mixture (3) was filtered to remove proteins, and then incubated with CcbS, PLP and L-glutamate. ($\star$ 1: GMP, $\star$ 2: residual **4-18**, #: a possible side product (**4-30**) due to over-reduction.)

under prolonged incubation. The recorded molecular mass of this product was 648.0960 (calcd. for $C_{18}H_{29}N_5O_{17}P_2^-$ $[M-H]^-$, 648.0960), which is consistent with that of reduced ketosugar after loss of one water molecule ($664 - 18 + 2 = 648$). Formation of this side product may be due to over-reduction during the dehydration reaction.

Because previous results showed that **4-18** is not a substrate for dehydration (Figure 4.8, trace 6), formation of a ketosugar product **4-25** different from **4-24** is possibly resulted from a different substrate. It is thus proposed that the intermediate of C-4 epimerization, the C-4 ketosugar **4-29**, could be recognized and processed by LmbL/CcbZ to undergo dehydration as well. The C-4 ketosugar intermediate (**4-29**) could be generated as the leaking product during the LmbM-catalyzed epimerization reaction (Figure 4.12B). The reaction mechanism of LmbM is expected to resemble that of UDP-galactose 4-epimerase (UGE), which is the most extensively studied 4-epimerase in the SDR family. The deprotonation from the C-4 hydroxyl group followed by the C-4 hydride transfer to NAD^+ leads to a UDP-4-ketohexose intermediate (**4-26** $\rightarrow$ **4-27**, Figure 4.12A). The re-orientation of the ketosugar intermediate (**4-27** $\rightarrow$ **4-27′**) and the re-addition of the C-4 hydride from NADH to the opposite side of the ketosugar results in the formation of the epimer **4-28**. It has been reported that the keto-intermediate **4-27** could be occasionally prematurely released. The amount of the release **4-27** accumulated upon extended incubation.[163] Therefore, it is likely that LmbM-catalyzed reaction produced two products, the C-4 epimer **4-19** (major) and the C-4 ketosugar **4-29** (minor), and both of them could serve as the substrates for dehydration.

Productive and non-productive ketosugars

As two different ketosugar products are generated from the LmbL/CcbZ-catalyzed dehydration under different conditions, it is important to examine whether

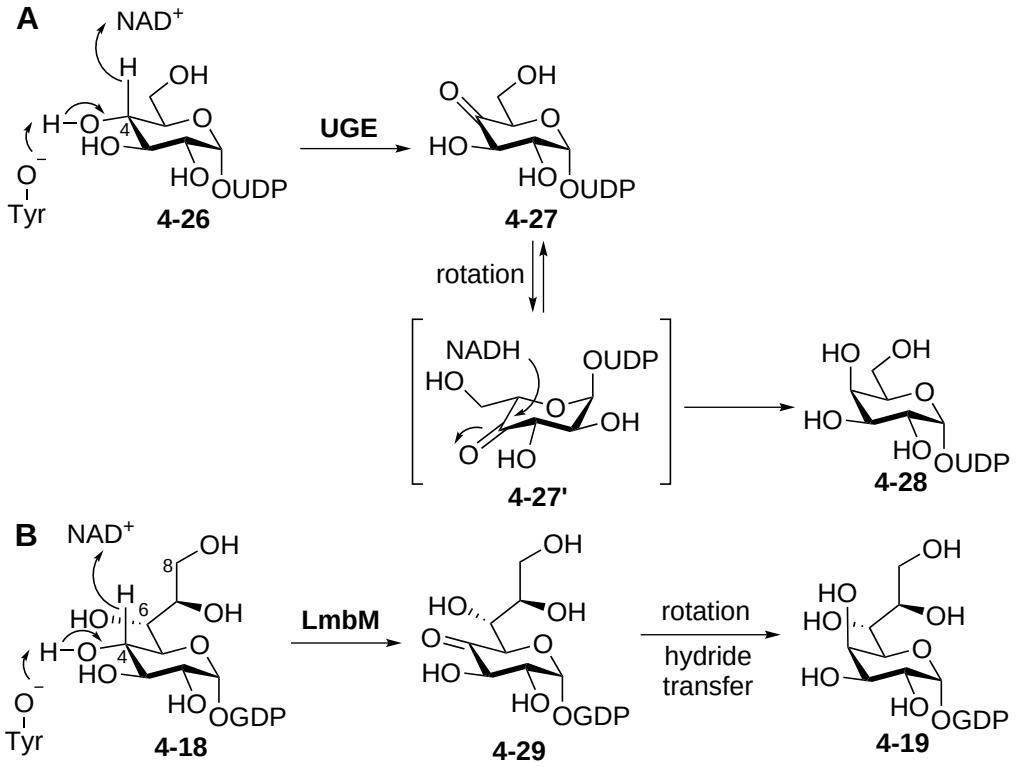


Figure 4.12: (A) Mechanism of UDP-galactose 4-epimerase. (B) Proposed mechanism of LmbM-catalyzed epimerization.

these ketosugars show chemical competency in the later enzymatic modifications. If only one ketosugar is the relevant biosynthetic intermediate in the lincomycin biosynthesis, the putative aminotrasnferase, CcbS, should be able to convert that specific GDP-ketosugar to the corresponding GDP-aminooctose. Coupled enzyme assays using **4-18** or **4-19** were performed to probe this hypothesis. After the formation of **4-24** and **4-25** was observed in the reaction mixtures, the enzymes were removed through filtration and the filtrate was subjected to transamination by adding CcbS and L-glutamate. In trace 4 of Figure 4.10, a new peak with retention time ~8 min was detected in the sample containing ketosugar **25**. The ESI-MS analysis indicated that the molecular weight matched the value predicted for a GDP-aminooctose (calcd. for $C_{18}H_{30}N_6O_{16}P_2^-$ $[M-H]^-$, 647.1121; obsd, 647.1124). The synthetic standard of the predicted GDP-aminooctose (**4-23**), which was prepared by another graduate student Richiro Ushimaru, showed consistent retention time on HPLC with the aminosugar generated from CcbS reaction (trace 5). In contrast, the ketosugar **4-24** generated from LmbL/CcbZ reaction using **4-19** as substrate could not be accepted by CcbS (trace 2). Therefore, the ketosugar **4-24** is likely a non-productive side product, while **4-25** serves as a productive biosynthetic intermediate.

Reverse transamination catalyzed by CcbS

CcbS was shown to catalyze the transamination of the ketosugar **4-25** to an aminooctose **4-23**. Because PLP-dependent enzymatic transamination is a reversible process, CcbS should be able to convert the synthetic aminooctose **4-23** to the corresponding ketosugar **4-25**.[164] To verify the function of CcbS, a reaction mixture of CcbS, α-ketoglutarate (α-KG), the synthetic aminooctose **4-23** and PLP prepared. A new product with a retention time ~18 min was detected by HPLC (Figure 4.13, trace 2 and 3). This peak has the same mass value as that of a dehy-

drated product, and its retention time is similar to that of the ketosugar **4-25** (trace 5 and 6) rather than **4-24** (trace 4). These results further support that the ketosugar **4-25** is the biosynthetically relevant intermediate in the GDP-octose modification pathway.

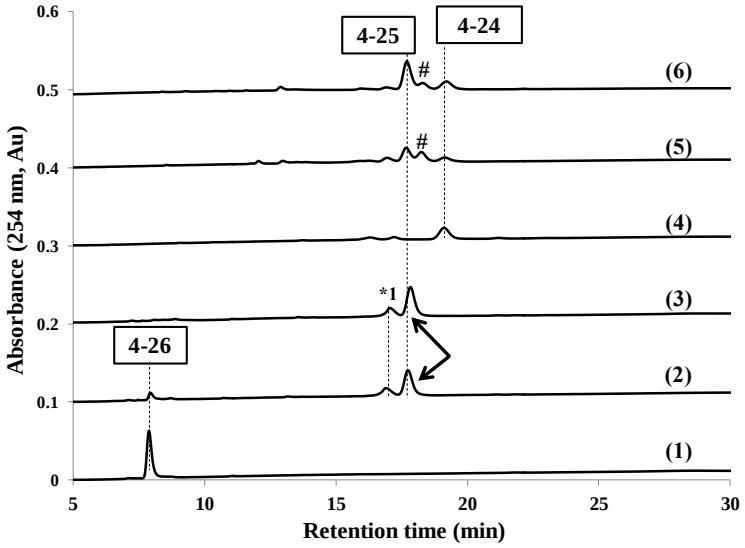


Figure 4.13: HPLC analysis of reverse transamination catalyzed by CcbS using the synthetic **4-23** as substrate. (1) Synthetic standard of **4-23**. (2) Aminosugar **4-23** incubated with CcbS, PLP and α-KG at 30 °C for 1 h. (3) Aminosugar **4-23** incubated with CcbS, PLP and α-KG at 30 °C for 3 h. (4) GDP-sugar **4-19** incubated with LmbL and CcbZ at 30 °C for 3 h. (5) GDP-sugar **4-19** incubated with LmbL, LmbM and CcbZ at 30 °C for 3 h. (6) GDP-sugar **4-18** incubated with LmbL, LmbM and CcbZ at 30 °C for 3 h. ($\star$ 1: impurity from α-KG, #: a possible side product (**4-30**) due to over-reduction.)

4.3.5 Mechanistic Proposal of the Epimerization/Dehydration Catalyzed by LmbL, LmbM and CcbZ

Based on the above-mentioned results, a mechanistic proposal regarding the epimerization and dehydration was proposed (Figure 4.14). LmbM was initially proposed as the C-4 epimerase in the lincomycin A biosynthetic pathway (Figure 4.12B) due to its homology to UDP-galactose epimerase, TDP-glucose 4,6-dehydratase and UDP-glucuronic acid decarboxylase in the SDR protein family. The prediction that LmbM has the potential to oxidize the C-4 hydroxyl group

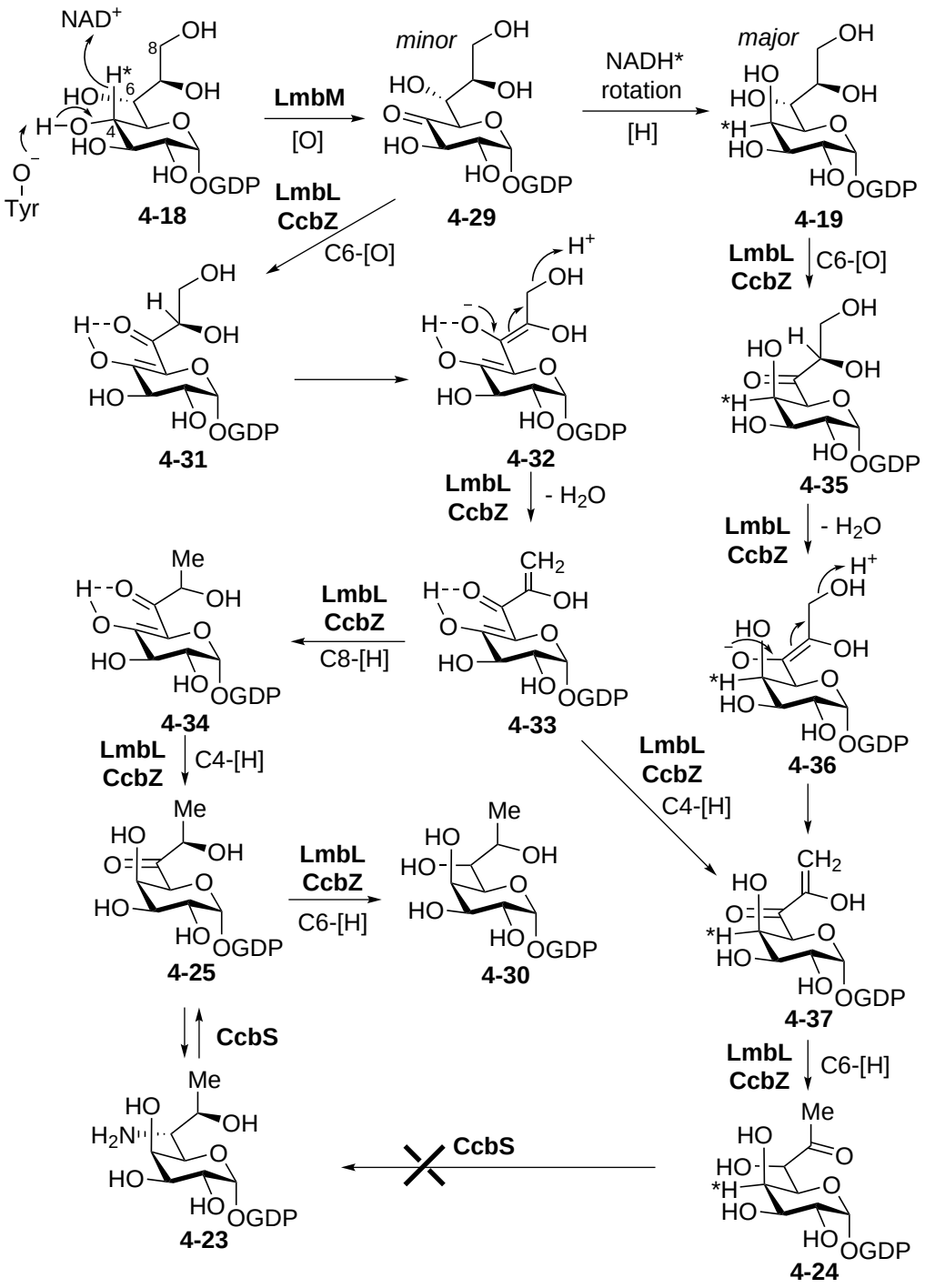


Figure 4.14: Proposed mechanism of GDP-octose **4-18** epimerization and dehydration.

combined with the biochemical characterization of the subsequent reactions revealed that the C-4 ketosugar **4-29** rather than the C-4 epimerized product **4-19** may actually be the relevant intermediate in the lincomycin A biosynthetic pathway.

The results from the enzymatic assays suggested that both LmbL and CcbZ are required for the formation of ketosugars (**4-24** and **4-25**). It is intriguing that the overall redox-neutral dehydration reaction requires two NAD^+-dependent enzymes. As discussed earlier (page 111), the dehydration is usually initiated with the dehydrogenation of the nearby hydroxyl group of a NDP-sugar substrate (**4-38** $\rightarrow$ **4-39**, Figure 4.15).[11,165] The resulting keto group, in turn, lowers the pKa of the α-proton facilitating α-deprotonation followed by β-elimination to expel a water molecule (**4-39** $\rightarrow$ **4-40**). The subsequent hydride transfer proceeds in a 1,4-addition manner to reduce the double bond and yields a ketosugar product (**4-40** $\rightarrow$ **4-41**). These three steps, dehydrogenation, elimination of water, and hydride transfer, are typically catalyzed by a single dehydratase and occur in the same active site. An internal hydride return from the transiently reduced cofactor to the α,β-unsaturated ketone intermediate is the key feature of the dehydration reaction. However, if multiple enzymes are involved in the dehydration mechanism, it is possible that these mechanistic steps may occur in the active-sites of different proteins.

Since the equilibrium of the first C-6 dehydrogenation, catalyzed by either LmbL or CcbZ, possibly favors the hydroxyl substrate, the amount of the C-6 ketosugar intermediates (**4-31** and **4-35**) may be too little to be detected in the reaction solution. It is also possible that the dehydrogenase binds tightly with its oxidization intermediate and the release of intermediate only occurs sporadically. Under normal catalytic condition, the elimination of water and/or hydride transfer would drive the overall dehydration reaction forward.

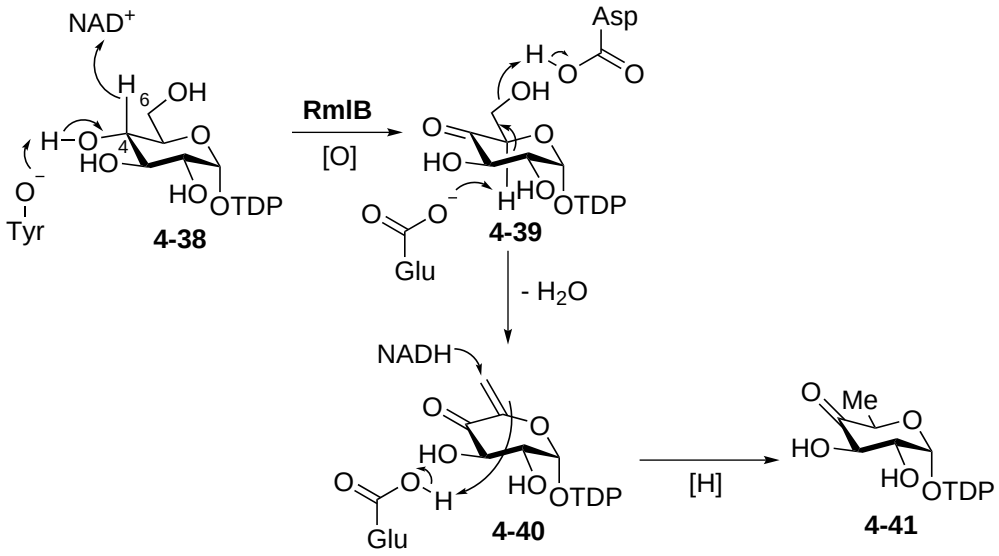


Figure 4.15: Mechanism of TDP-glucose 4,6-dehydratase.

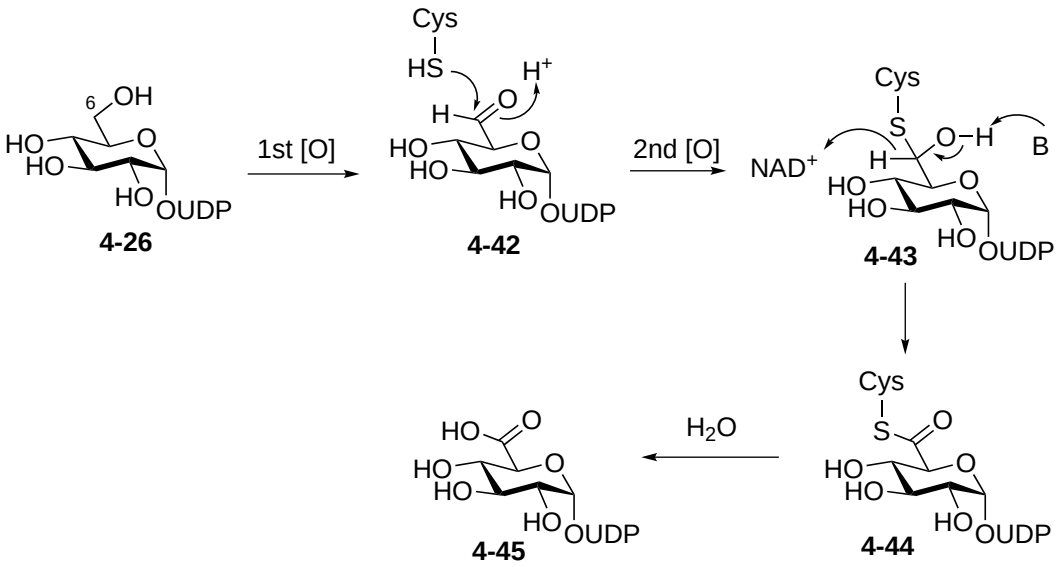


Figure 4.16: Mechanism of UDP-glucose 6-dehydrogenase.

Although the actual catalyst for each mechanistic step remains unclear based on the current data, LmbL is proposed as the C-6 dehydrogenase based on the amino acid sequence analysis. LmbL was annotated as a putative UDP-glucose 6-dehydrogenase, which catalyzes the oxidation of UDP-glutose to UDP-gluconate (**4-26** $\rightarrow$ **4-45**, Figure 4.16). In the reaction mechanism for UDP-6-glucose dehydrogenase, a sequestered C-6 aldehyde intermediate **4-42** is produced in the first oxidation step and a covalently bound thiohemiacetal intermediate (**4-43**) is the substrate for the second oxidation reaction (Figure 4.16).[166] The conserved L-cysteine residue, Cys260, in the active site presumably acts as the nucleophile to trap the aldehyde. Although LmbL is believed to be responsible for oxidizing the C-6 hydroxyl group of a NDP-sugar, a close examination of the primary sequence of LmbL reveals a less nucleophilic amino acid, L-threonine, at the site of the conserved L-cysteine residue in the active site. Interestingly, a C260A mutant of UDP-glucose 6-dehydrogenase, which is similar to LmbL (C261T), has been constructed and reported in the literature. Mechanistic studies using the C260A mutant showed that it is capable of both reducing the aldehyde intermediate **4-42** (with NADH), but is incapable of oxidizing UDP-glucose to UDP-gluconate (with NAD^+). It was thus reasoned that the C260A mutant might form a ternary complex with the aldehyde intermediate and NADH. The oxidation of aldehyde to carboxylic acid does not proceed and the release of the bound aldehyde occurs slowly.[166] This provides a probable explanation for the fact that ketosugar intermediates (**4-31** and **4-35**) were not detected in the dehydration reaction.

Out current data did not allow a precise prediction of which enzyme is responsible for water elimination and the regiochemistry of hydride transfer. CcbZ is proposed to catalyze the downstream reductions via hydride transfer from NADH. Since the closest homolog of Ccbz with known function is *myo*-inositol dehydrogenase,[167] the substrate binding of Ccbz may not rely on the recognition of the

nucleotide-diphosphate moiety. Instead, CcbZ may possess a relatively flexible substrate specificity and accepts various poly-hydroxylated molecules as substrates. The formation of the over-reduced side-product (**4-30**) might be resulted from the structural tolerance of the final reductase.

4.4 Conclusion

In conclusion, we have shown the formation of GDP-aminooctose (**4-23**) by combining LmbL, LmbM, CcbS and CcbZ with the GDP-octose substrate (**4-18**). Although the initial characterization of LmbM as the C-4 epimerase suggested the GDP-octose modification follows the route A in Figure 4.4, the later observation of one productive ketosugar intermediate **4-25** and the other non-productive ketosugar side product **4-24** led to a modified pathway shown in Figure 4.14. The complicated interplay between these three SDR enzymes remains unclear. One hypothesis is that the keto group-containing compounds are labile under physiological conditions, and thus the tight-binding of the intermediates and the grouping of the operating enzymes may prevent the ketosugars from leaking out of the biosynthetic machinery. It is possible that there are physical interactions between LmbL and CcbZ, so that two proteins catalyze the dehydration reaction as a complex. Nevertheless, a mechanistic proposal which incorporates results from biochemical characterization and sequence analysis is proposed. This study represents the first characterization of GDP-octose modifications including C-4 epimerization, C-6 amination and C-8 dehydration and sets the stage for studying the unusual octose biosynthetic pathways. This work also has practical implications as it provides the enzyme tools for manipulating high-carbon chain-containing NDP-sugars. Moreover, the GDP-aminooctose (**4-23**) is expected to be the final product of the serial GDP-octose modifications and the crucial substrate for C-1 sulfur incorporation and N-6 amide bond formation.

Chapter 5

Biosynthetic Studies of Lincomycin A (IV): Sulfur Incorporation and Amide Bond Formation

5.1 Introduction

In the last chapter, biosyonthetic processing of GDP-aminooctose intermediate was successfully demonstrated *in vitro* and revealed a complicated interplay among three GDP-sugar modifying enzymes. The identified GDP-aminoocotse not only represents a novel type of NDP-sugar, but also serves as the immediate precursor for the sulfur transfer reaction in lincomycin A biosynthesis. The GDP-aminooctose is certainly an important intermediate because its structure contains all chemical features required for sulfur incorporation and amide bond formation. The C-1 GDP moiety is clearly a good leaving group for the substitution by a thiol nucleophile to form a C–S bond. Likewise, the C-6 amino group may engage in the C–N bond formation through a reaction involving an activated carboxylate.

The known examples of sulfur incorporation have been divided into two types based on their C–S bond formation mechanisms.[168] In the first type, the sulfur incorporation proceeds through an ionic mechanism involving nucleophilic attack of a sulfur donor to an electrophilic acceptor. Various sulfur-containing molecules, including inorganic bisulfide (HS^-), thiocarboxylate group ($R-COS^-$), cysteine persulfide group ($R-S-S^-$), and low-molecular-weight (LMW) thiols, have been shown to be the relevant sulfur source. The second type involving radical mechanism is usually catalyzed by radical SAM-enzymes containing multiple iron-sulfur clusters. The incorporated sulfur atom is either abstracted from the

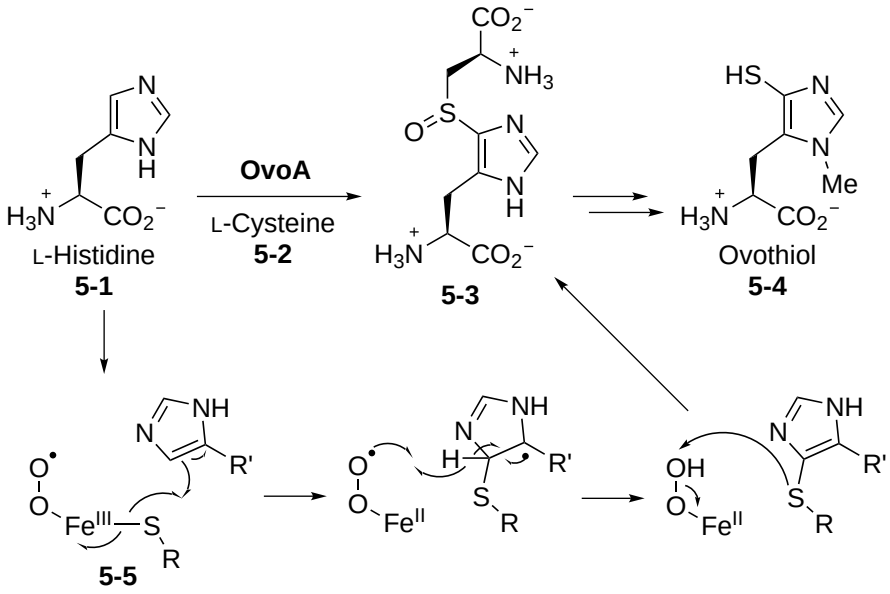


Figure 5.1: Proposed mechanism for OvoA-catalyzed oxidative sulfur insertion.

metal clusters or the protein-bound persulfides. One well-documented example involving radical mechanism is the thiocycle formation of biotin catalyzed by BioB (Section 1.1.2.1, page 4). Another interesting example of radical sulfur insertion was recently observed in the biosynthesis of ovothiol (**5-4**).[169,170] The C–S bond formation between the imidazole moiety of L-histidine (**5-1**) and the sulfur atom of L-cysteine (**5-2**) is catalyzed by a non-heme iron enzyme, OvoA. It has been proposed that the iron(III)-superoxide likely generates a thiyl radical (**5-5**) which subsequently attacks the imidazole ring (Figure 5.1).

The inorganic bisulfide ion, derived from the bacterial sulfur assimilation pathway, is the simplest sulfur source used in the ionic sulfur insertion reaction. For example, the PLP-dependent enzyme, *O*-acetylserine (**5-8**) sulfhydrylase (OASS), can utilize bisulfide in L-cysteine (**5-2**) biosynthesis (Figure 5.2A).[171] However, for most of the ionic sulfur insertions, the actual sulfur source are protein-bound persulfides (**5-10**) generated by cysteine desulfurases (Figure 5.2B).[172] This class of enzymes utilize the cofactor PLP (**5-6**) to cleave the C–S bond in the L-cysteine

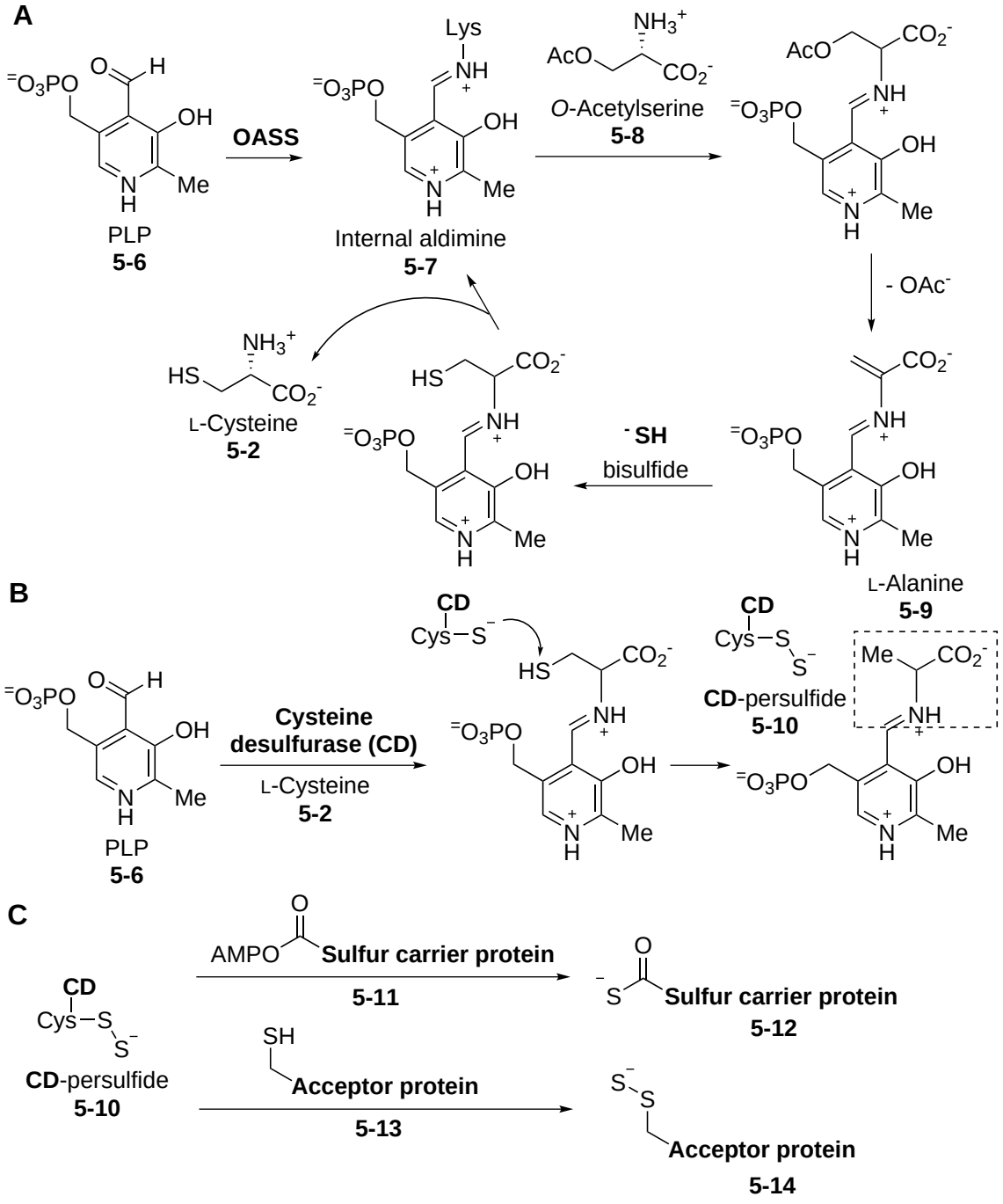


Figure 5.2: Ionic sulfur insertion mechanisms. (A) Cysteine biosynthesis catalyzed by OASS. (B) Formation of protein persulfide catalyzed by cysteine desulfurase (CD). (C) Formation of acceptor protein persulfide and protein thiocarboxylate.

(**5-2**) and produce L-alanine (**5-9**) and sulfane sulfur (S^0), the latter of which is subsequently coupled to an active site cysteine residue of the desulfurase to form a persulfide group (**5-10**). The terminal sulfur of the resulting persulfide is then transferred to cysteine residues in an acceptor protein to make a new persulfide group (**5-14**). In turn, the reactive terminal sulfur atom of the resulting persulfide on the acceptor protein can act as a nucleotphile to attack an electrophilic precursor to form a C–S bond, which is the mode of sulfur incorporation observed in thionucleoside biosynthesis.[173] The protein-bound persulfide can also attack an activated carboxylate group (**5-11**) to form a thiocarboxylate group at the *C*-terminus of a sulfur carrier proteins (**5-12**). Transfer of sulfur stom from a thiocarboxylate group has been observed in thiamine and BE-7585A biosynthesis (Section 1.1.2.2, page 7).

Low-molecular-weight (LMW) thiols shown in Figure 5.3, are mostly considered as thiol-redox buffers to maintain intracellular redox balance and as post-translational thiolation reagents to regulate protein functions.[174] Occasionally, these thiols are found to mediate the detoxification of different reactive species in xenobiotic metabolism[174] or chaperon the metal cluster delivery in the cytoplasm.[175] It is uncommon to see LMW thiols directly involved in the biosynthesis of natural products via the ionic sulfur insertion. Nevertheless, in the biosynthesis of glucosinolates, cysteine or glutathione (GSH, **5-15**) have been suggested as the sulfur source (Section 1.1.2.5, page 10). However, the detailed mechanism and the identity of the enzymes catalying the sulfur transfer remain elusive. Recent investigation of the epidithiodioxopiperazine (ETP) biosynthesis demonstrated that an unusual glutathione *S*-transferase (GST), GliG, is able to catalyze the C–S bond formation between GSH and a bishydroxylated piperazine precursor.[176,177] So far, this is the only established example indicating the involvement of a LMW thiol, GSH, in enzymatic C–S bond formation in the biosynthesis of secondary metabolites.

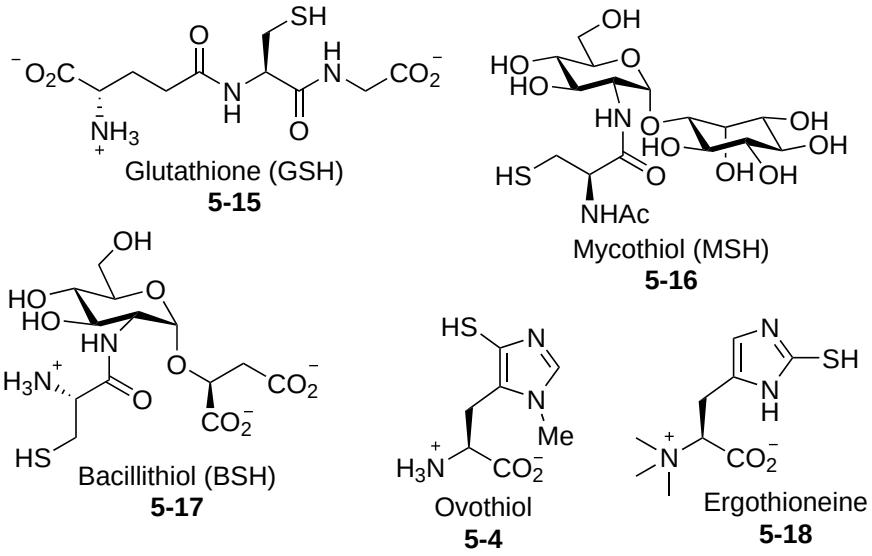


Figure 5.3: Structures of bacterial low molecular weight thiols.

Combing the knowledge from the literature and the results of sequence analysis of the *lmb* gene cluster discussed in Section 2.3.1, we proposed that mycothiol (MSH, **5-16**), which also belongs to LMW thiols, may serve as the sulfur donor in lincomycin A biosynthesis. The biosynthetic pathway describing the sulfur transfer, amide bond formation and post-sulfur transfer modifications is proposed in Figure 5.4. A putative MSH-*S*-transferase, LmbV, likely recognizes and enhances the nucleophilicity of MSH, allowing the C–S bond formation occur at the C-1 position of GDP-aminoocotose (**5-19**). The electrophilicity of GDP-sugar may be further increased by the action of LmbT, a putative glycosyltransferase (GT). This class of enzymes are known to catalyze the reversible glycosidic bond formation via the intermediacy of an oxocarbenium ion.[178] Furthermore, considering that the GDP-aminooctose (**5-19**) also needs to undergo amide bond formation at C-6 position, the amide bond may be formed prior to (**5-19** $\rightarrow$ **5-24**) or after the sulfur transfer (**5-20** $\rightarrow$ **5-25**). In fact, it may occur at maby possible stages in the pathway (see Figure 5.4). Hydrolysis of the MSH-sugar adducts (**5-20** and **5-25**) is expected to be catalyzed by LmbE, a protein homolog of MSH *S*-conjugate ami-

dase. LmbF is proposed as the PLP-dependent lyase responsible for cleaving the C–S bond to release the free thiol intermediates (**5-22** and **5-27**). Anomers of **5-22** or **5-27** may interconvert to each other through the open chain form until the SAM-dependent methylation (catalyzed by LmbG) leading to final product having an α-configuration at C–1 position.

5.2 Experimental Procedures

5.2.1 General

Materials

All chemicals and reagents were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA) and were used without further purification unless otherwise specified. Lincomycin hydrochloride was purchased from Chem-Impex International, Inc (Wood Dale, IL). Oligonucleotide primers were prepared by Integrated DNA Technologies (Coralville, IA). Kits for DNA gel extraction and spin minipreps were purchased from Qiagen (Valencia, CA). PureLink® Genomic DNA Mini Kit was purchased from from Invitrogen (Carlsbad, CA). KOD NDA polymerase was purchased from Novagen (Madison, WI). QickChange site-directed mutagenesis kit was obtained from Stratagene, which was acquired by Agilent (Santa Clara, CA). Enzymes and molecular weight standards used for the cloning experiments were obtained from New England Biolabs (Ipswich, MA). Reagents for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) were purchased from Bio-Rad (Hercules, CA), with the exception of the protein molecular weight markers, which were obtained from Invitrogen. Growth medium components were acquired from Becton Dickinson (Sparks, MD). Sterile syringe filters are products of Fisher Scientific. Thiopropyl sepharose 6B was purchased from GE Healthcare Life Sciences (Piscataway, NJ). Amicon YM-10 ultrafiltration membranes were bought from Millipore (Billerica,

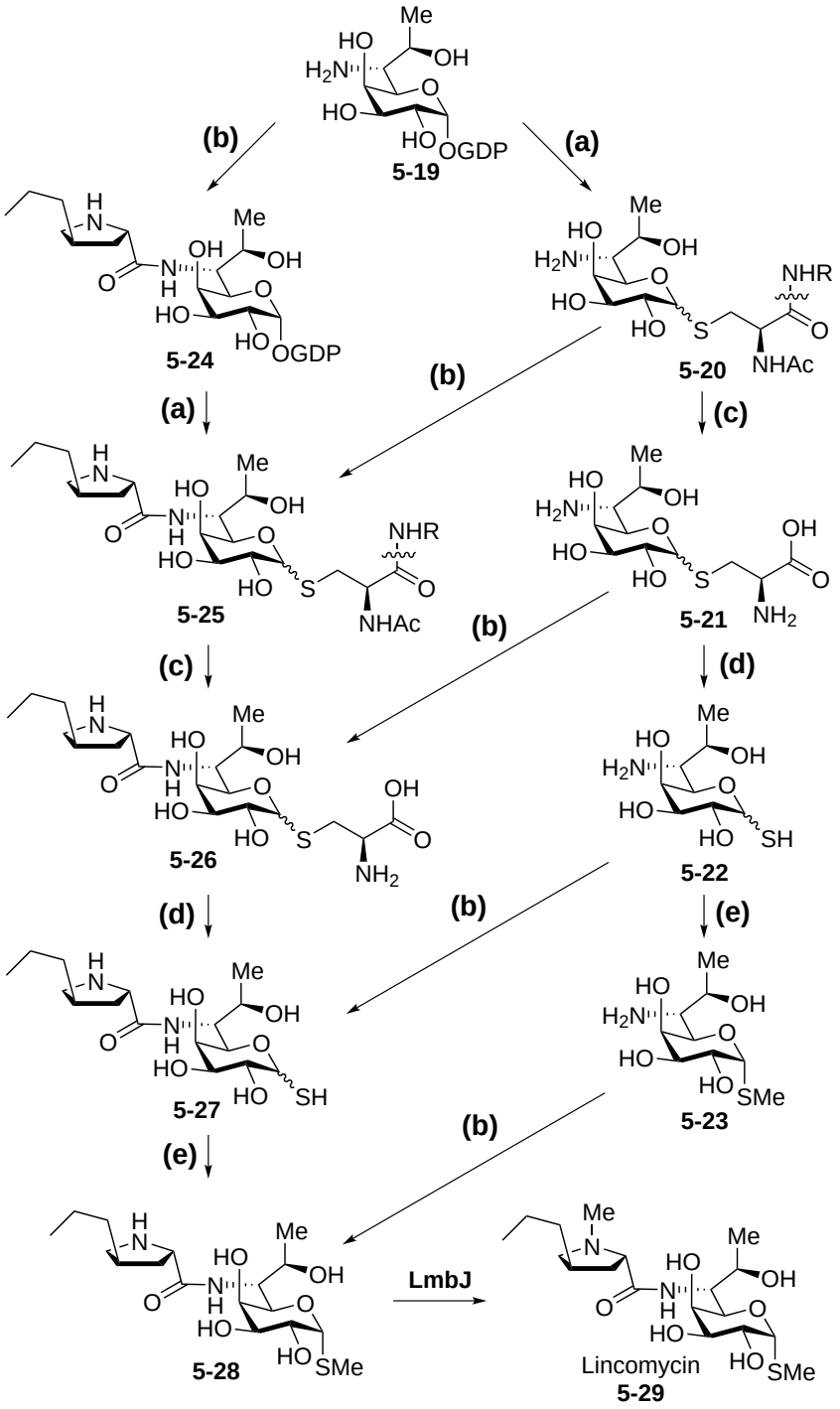


Figure 5.4: Later stage of the proposed lincomycin A biosynthetic pathway.

MA). The analytical and semi-preparative CarboPac PA1 high-performance liquid chromatography (HPLC) columns were obtained from Dionex (Sunnyvale, CA). Analytical C-18 HPLC columns are products of Varian (Palo Alto, CA). Simipreparative C-18 HPLC columns were purchased from Fisher Scientific (Pittsburgh, PA).

Bacterial Strains and Plasmids

Streptomyces lincolnensis NRRL ISP-5355 (identical to ATCC 25466) and *Streptomyces caelestis* NRRL 2418 were obtained from the Agricultural Research Service (ARS) Culture Collection of the National Center for Agricultural Utilization Research (Peoria IL). *Escherichia coli* DH5α, acquired from Bethesda Research Laboratories (Gaithersburg, MD), was used for routine cloning experiments. The protein overexpression host *E. coli* BL21 star (DE3) was obtained from Invitrogen. Vector pET28b(+) for protein overexpression was purchased from Novagen (Madison, WI).

Instrumentation

Standard genetic manipulations of *E. coli* were performed as described by Sambrook and Russell.[142] DNA sequencing was performed at the core facility of the Institute of Cellular and Molecular Biology, The University of Texas at Austin. Vector NTI Advance 10.1.1 from Invitrogen was used for sequence alignments. DNA concentrations were measured using a NanoDrop ND-1000 UVvis instrument from Thermo Fisher Scientific. HPLC was performed on a Beckman Coulter System Gold HPLC equipped with a UV detector or a Corona CAD (charged aerosol detector, ESA Biosciences, Chelmsford, MA). Mass spectra were recorded at the Mass Spectrometry core facility in the Department of Chemistry and the Proteomics Facility in College of Pharmacy at the University of Texas, Austin.

Table 5.1: Primers used for constructing plasmids containing genes involved in the sulfur transfer and post-sulfur transfer modifications.

Plasmid	Primer pair
pET24b(+)-*lmbC*-CHis_6	CIL030 / CIL031
pET24b(+)-*lmbD*-CHis_6	CIL032 / CIL033
pET28b(+)-*lmbD*-NHis_6	CIL032 / CIL034
pET28b(+)-*lmbE*-NHis_6	CIL035 / CIL036
pET28b(+)-*lmbF*-NHis_6	CIL037 / CIL038
pET28b(+)-*lmbG*-NHis_6	CIL039 / CIL040
pET24b(+)-*lmbT*-CHis_6	CIL041 / CIL042
pET28b(+)-*lmbT*-NHis_6	CIL041 / CIL043
pET24b(+)-*lmbV*-CHis_6	CIL044 / CIL045
pET28b(+)-*lmbV*-NHis_6	CIL044 / CIL046
pET28b(+)-*ccbD*-NHis_6	CIL047 / CIL048
pET28b(+)-*ccbF*-NHis_6	CIL049 / CIL050
pET28b(+)-*ccbT*-NHis_6	CIL051 / CIL052
pET28b(+)-*ccbV*-NHis_6	CIL053 / CIL054

Primer number	Primer name	Sequence[a]
CIL030	F-*lmbC*-24/28-NdeI	5′-TCTATAA**CATATG**TCGTCCTCCGTTCGACTCTCC-3′
CIL031	R-*lmbC*-24-HindIII	5′-ATAT**AAGCTT**CTCCCCGCGTGTGACG-3′
CIL032	F-*lmbD*-24/28-NdeI	5′-TCTATAA**CATATG**GCGGTATCACCCCAGTCG-3′
CIL033	R-*lmbD*-24-HindIII	5′-ATAT**AAGCTT**CAGGACCTCCGCGAGC-3′
CIL034	R-*lmbD*-28-HindIII	5′-TTAC**AAGCTT***CA*CAGGACCTCCGCGAGC-3′
CIL035	F-*lmbE*-24/28-NdeI	5′-TCTATAA**CATATG**ACTCAGTGCCTGCTGACC-3′
CIL036	R-*lmbE*-28-HindIII	5′-TTAC**AAGCTT***CA*CGCGGGAGCGG-3′
CIL037	F-*lmbF*-24/28-NdeI	5′-TACTAAC**CATATG**ACCGCCACGGCGAGC-3′
CIL038	R-*lmbF*-28-HindIII	5′-ATAT**AAGCTT***CA*CCGGTACCGCCACTCG-3′
CIL039	F-*lmbG*-24/28-NdeI	5′-TACTAAC**CATATG**CGGGACTACCGTCTCTGGC-3′
CIL040	R-*lmbG*-28-HindIII	5′-ATAT**AAGCTT**GGA*TCA*GCCAGCCGTCGTCG-3′
CIL041	F-*lmbT*-24/28-NdeI	5′-AGTTAAG**CATATG**ACGGCGCGGACGG-3′
CIL042	R-*lmbT*-24-HindIII	5′-GGAT**AAGCTT**TGACACCTCCGCCAGC-3′
CIL043	R-*lmbT*-28-HindIII	5′-GTCG**AAGCTT***CA*TGACACCTCCGCCAGC-3′
CIL044	F-*lmbV*-24/28-NdeI	5′-AGTTAAG**CATATG**CAGCGCAAGGGACTGG-3′
CIL045	R-*lmbV*-24-HindIII	5′-ATAT**AAGCTT**CGGCCCCACCAGCACC-3′
CIL046	R-*lmbV*-28-HindIII	5′-GTAA**AAGCTT***CA*CGGCCCCACCAGC-3′
CIL047	F-*ccbD*-24/28-NdeI	5′-TCTATAA**CATATG**GCCCAATCCAAGGGTTCG-3′
CIL048	R-*ccbD*-28-HindIII	5′-TTAC**AAGCTT***CA*GAGTTCCTTGAGCAATCG-3′
CIL049	F-*ccbF*-24/28-NdeI	5′-TCTATAA**CATATG**TCCGACTTAGCTGCCG-3′
CIL050	R-*ccbF*-28-HindIII	5′-TTAC**AAGCTT***CA*GCGGGGCTGC-3′
CIL051	F-*ccbT*-24/28-NdeI	5′-TCTATAA**CATATG**GCTGAAGCGGCTTCG-3′
CIL052	R-*ccbT*-28-HindIII	5′-TTAC**AAGCTT***CA*CAGGTCGCTCCAGC-3′
CIL053	F-*ccbV*-24/28-NdeI	5′-TCTATAA**CATATG**GAAGGGTTACGCCGACTGG-3′
CIL054	R-*ccbV*-28-HindIII	5′-TTAC**AAGCTT**GGC*TCA*CAGGAGAACTTTGAG-3′

[a] The engineered restriction sites are shown in bold; the start codon is shown in bold and underlined; the stop codon is shown in italic.

5.2.2 Cloning of Genes Involved in the Sulfur Transfer, Amide Bond Formaion and Post-sulfur Transfer Modifications.

Plasmids containing *lmbC*, *lmbD*, *lmbE*, *lmbF*, *lmbG*, *lmbT*, *lmbV*, *ccbD*, *ccbF*, *ccbT* and *ccbV* were constructed by amplifying the corresponding genes from genomic DNA isolated from *S. lincolnensis* and *S. caelestic* using designed primer pairs (Table 5.1). The PCR products were then digested with NdeI and HindIII and ligated into pET24b(+) and pET28b(+), which had been digested with the same restriction enzymes. The resulting plasmids were sequenced using T7 or T7 terminal universal primer and used to transform the *E. coli* BL21 star (DE3) for protein overexpression.

Table 5.2: Primers used for constructing plasmids containing genes encoding thiolation domains within *lmbN* and *ccbZ*.

Plasmid	Primer pair
pET28b(+)-*lmbN-T-N*His_6	CIL055 / CIL056
pET28b(+)-*ccbZ-T-N*His_6	CIL057 / CIL058

Primer number	Primer name	Sequence[a]
CIL055	F-*lmbN-T*-24/28-NdeI	5′-ATCACAA**CATATG**AGCACTCTGGACGAGGTCC-3′
CIL056	R-*lmbN-T*-28-HindIII	5′-TTAC**AAGCTT***CA*ATGGGTGATACGGGCCACC-3′
CIL057	F-*ccbZ-T*-24/28-NdeI	5′-TCTATAA**CATATG**CCGTCGCTGCTGGTCG-3′
CIL058	R-*ccbZ-T*-28-HindIII	5′-TTAC**AAGCTT***CA*CGTGCTGGTGCGG-3′

[a]The engineered restriction sites are shown in bold; the start codon is shown in bold and underlined; the stop codon is shown in italic.

5.2.3 Cloning of Thiolation Domains Encoded in *lmbN* and *ccbZ*

Plasmids containing genes encoding the putative thiolation domains within *lmbN* and *ccbZ* were constructed by amplifying the corresponding fragments from pET24b(+)-*lmbN*-CHis_6 and pET28b(+)-*ccbZ-N*His_6, respectively, using designed primer pairs (Table 5.2). The PCR products were then digested with NdeI and

HindIII and ligated into pET28b(+), which had been digested with the same restriction enzymes. The resulting plasmids were sequenced using T7 or T7 terminal universal primer, and used to transform the *E. coli* BL21 star (DE3) for protein overexpression.

5.2.4 Overexpression and Purification of His-tagged Proteins in *E. coli*

Overexpression and purification procedures for *C*-His$_6$-LmbC, *N*-His$_6$-LmbE, *N*-His$_6$-LmbF, *N*-His$_6$-LmbG, *N*-His$_6$-CcbD, *N*-His$_6$-CcbF, *N*-His$_6$-CcbT and *N*-His$_6$-CcbV were identical to those used to purify *N*-His$_6$-LmbK as described in Sections 3.2.4 and 3.2.5. The yields of the *N*-His$_6$ tagged proteins from 1 L culture were ∼17 mg for *C*-His$_6$-LmbC, ∼55 mg for *N*-His$_6$-LmbE, ∼10 mg for *N*-His$_6$-LmbF, ∼20 mg for *N*-His$_6$-LmbG, ∼5 mg for *N*-His$_6$-LmbN-T, ∼14 mg for *N*-His$_6$-CcbD, ∼9 mg for *N*-His$_6$-CcbF, ∼9 mg for *N*-His$_6$-CcbT, ∼8 mg for *N*-His$_6$-CcbV and ∼1 mg for *N*-His$_6$-CcbZ-T. The molecular mass and purity of the purified proteins were determined by SDS-PAGE analysis.

5.2.5 Preparations of Methylthiolincosamide (5-23)

Methylthiolincosamide (5-23).

Compound **5-23** was obtained from the commercial available lincomycin A (**5-29**) hydrochloride salt following a previously reported procedure.[179] Lincomycin A (0.75 g, 1.69 mmole) was added hydrazine monohydrate (7.5 mL) and heated to reflux for 24 h. After cooling down the reaaction mixture, the volatiles were removed by rotary evaporation under reduced pressure. The residues were then treated with acetonitrile and stirred until all of the lumps had broken up. The white solids were collected and washed with acetonitrile and then ether. This material was then subjected to recrystallization from methanol and water. The crystals separated from the solution were dried to yield methylthiolincosamide (**5-23**)

($C_9H_{19}NO_5S$, 0.42 g, 98%) as white solids. ^{1}H NMR (D_2O, 400 MHz) δ (ppm) 5.29 (1 H, d, $J_{1,2}$ = 6.4 Hz, H1), 4.21 (1 H, dd, $J_{4,5}$ = 7.2 Hz, $J_{5,6}$ = 0.8 Hz, H5), 4.10 (1 H, dd, $J_{1,2}$ = 6.4 Hz, $J_{2,3}$ = 10.0 Hz, H2), 4.10–4.08 (1 H, m, H6), 4.05 (1 H, dq, $J_{6,7}$ = 0.4 Hz, $J_{7,8}$ = 6.4 Hz, H7), 3.57 (1 H, dd, $J_{2,3}$ = 10.0 Hz, $J_{3,4}$ = 3.2 Hz, H3), 3.25 (1 H, dd, $J_{4,5}$ = 7.2 Hz, $J_{3,4}$ = 3.2 Hz, H4), 2.09 (3 H, s, SME), 1.22 (3 H, d, $J_{7,8}$ = 6.4 Hz, H8). The recorded data is consistent with the values reported in literature.[180]

5.2.6 Purification of Mycothiol (5-16) from *S. lividans*

The spores of *S. lividans* 1326 were striped on 2% agar containing 1% glycerol, 0.5% yeast extract and 0.5% malt extract. The plates were maintained at 30 °C for 2 days. Seeding cultures (7 mL) were inoculated by a piece of agar containing colonies and incubated at 30 °C for 48 h. Tryptic soy broth (3% TSB and 1% glycerol) was chosen as the liquid medium since it contains very small amounts of GSH. Bulk cultures (500 mL) were inoculated with 1% of the starting culture in TSB medium without glycerol and shaken in 2-L flask at 200 rpm and 28 °C for 48 h or 72 h. Cells were then harvested and frozen at −80 °C until lysis.

The method for MSH preparation was modified from the reported procedure.[181] The pelleted cells (∼10 g wet weight) were suspended in 60 mL of pre-warmed lysis buffer (60 °C, 50% v/v aqueous acetonitrile containing 25 mM methanesulfonic acid) and then sonicated. The lysate was incubated at 60 °C for an additional 15 min. Cellular debris was removed by centrifugation, then the CH_3CN was evaporated by rotavap from the supernatant. DTT was added to reach a final concentration of 1 mM. Solid Tris · base was added with stirring until the pH reached 8. The supernatant was concentrated to 10 mL and kept at −20 °C until thiol-affinity column purification.

Thiol-affinity chromatography was performed by using Thiopropyl sepharose 6B. Following the protocol from the manufacturer and reported procedure,[181] dry

resins (2 g) were first hydrated with ddH_2O (30 mL) for 30 min and washed by ddH_2O (30 mL)×2. The wet medium reached a volume of ∼9 mL. The resins were equilibrated with 50 mL of 50 mM Tris pH 8, and then incubated with the concentrated cell-free extracts (from 10 g of cells, around 10 mL) for 2 h. After loaded, the column was rinsed first with 225 mL of 50 mM Tris pH 8 and then 100 mL of 50 mM NH_4HCO_3 pH 8. The bound thiols were finally eluted with 30 mL of 20 mM DTT in 50 mM NH_4HCO_3 pH 8. The collected eluent (30 mL) was lyophilized and the residues were redissolved in 10 mL of H_2O. MSH was purified by HPLC on an Econosil C-18 semiprep column, 10 μm, 250 × 10 mm (Alltech), with the following solvents: mobile phase A was 0.1% TFA in water and mobile phase B was methanol. The separation was achieved at a flow rate of 3 mL/min with a gradient program: 0-25 min 0% B, 25-27 min 0-10% B, 27-37 min 10% B, 37-39 min 10-0% B, 39-49 min 0% B. Mycothiol was eluted at the retention time of ∼10 min. Purified MSH was lyophilized and stored as a concentrated solution (∼920 μM, determined by DTNB assay) in 0.1% TFA. Overall yield of MSH from 10 g of cells is 1.87 mg.

5.2.7 Sulfur Transfer Reaction

The synthetic GDP-aminooctose (**5-19**, 120 μM) was incubated with *N*-His_6-CcbT protein (16 μM), a selected LMW thiol reagent (1.5 mM), dithiothreitol (DTT, 0.15 mM) in 30 mM Tris buffer (pH 8.0) at ambient temperature for 6 h. Other reactions containing only *N*-His_6-CcbV protein (16 μM) or both enzymes were similarly prepared. The LMW thiols tested included L-cysteine (**5-2**), GSH (**5-15**) and MSH (**5-16**). The control reaction omitting enzymes included all three thiols. At the end of incubation, each reaction mixture was filtered through a YM-10 membrane to remove proteins. HPLC analysis of the filtrate was performed using a Dionex CarboPac PA1 analytical column (4 × 250 mm) with the following solvents: mobile phase A was water and mobile phase B was 1 M NH_4OAc. The separation

was conducted at a flow rate of 1 mL/min with a gradient program: 0-2 min 10% B, 2-18 min 10-90% B, 18-22 min 90% B, 22-24 min 90-10% B, 24-30 min 10%B. The detector wavelength was set at 254 nm.

5.2.8 Functional Characterization of the Lincosamide Synthetase

Amide bond formation using MTL (5-23) or GDP-aminooctose (5-19) as substrates

The purified MTL (**5-23**, 1.25 mM) or GDP-aminoocotse (**5-19**, 150 μM) was incubated with a selected amino acids (3 mM), ATP (2 mM), $MgCl_2$ (12.5 mM), coenzyma A (1 mM), Sfp (8 μM), thermostable inorganic pyrophosphatas (2 units), *C*-His_6-LmbN (2 μM), *C*-His_6-LmbC (8 μM) and *N*-His_6-CcbD (8 μM) in 15 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (pH 7.0) at ambient temperature for 6 h. Amino acids tested included propylproline (**5-38**), L-proline, L-alanine, fusaric acid (**5-39**) and picolinic acid (**5-40**). The filtrates after the removal of proteins were subjected to LC-MS analysis.

Functional characterization of LmbN-T and CcbZ-T

The purified MTL (**5-23**, 1.25 mM) was incubated with propylproline (**5-38**, 3 mM), ATP (2 mM), $MgCl_2$ (12.5 mM), coenzyme A (1 mM), Sfp (8 μM), thermostable inorganic pyrophosphatas (2 units), *N*-His_6-LmbN-T or *N*-His_6-CcbZ-T (2 μM), *C*-His_6-LmbC (8 μM) and *N*-His_6-CcbD (8 μM) in 15 mM HEPES buffer (pH 7.0) at ambient temperature for 6 h. The filtrates after the removal of proteins were subjected to LC-MS analysis.

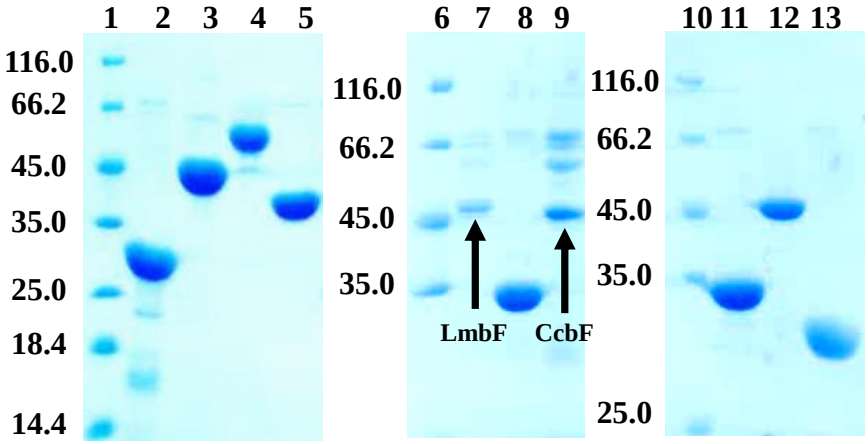


Figure 5.5: SDS-PAGE gel of the purified proteins: *C*-His$_6$-LmbN (47.4 kDa, lane 2), *N*-His$_6$-CcbZ (37.0 kDa, lane 3), *C*-His$_6$-LmbC (37.0 kDa, lane 4), *N*-His$_6$-CcbD (37.0 kDa, lane 5), *N*-His$_6$-LmbF (37.0 kDa, lane 7), *N*-His$_6$-LmbG (37.0 kDa, lane 8), *N*-His$_6$-CcbF (37.0 kDa, lane 9), *N*-His$_6$-LmbE (37.0 kDa, lane 11), *N*-His$_6$-CcbT (37.0 kDa, lane 12) and *N*-His$_6$-CcbV (37.0 kDa, lane 13). The molecular markers are shown in lanes 1, 6 and 10. Minor impurities were possibly resulted from the non-specific interactions with Ni-NTA resins.

5.3 Results and Discussion

5.3.1 Overexpression and Purification of His-tagged Proteins in *E. coli*

To investigate the sulfur incorporation and the amide bond formation in lincomycin A biosynthesis, the encoding genes for the enzymes involved, *lmbC*, *lmbD*, *lmbE*, *lmbF*, *lmbG*, *lmbT* and *lmbV*, were amplified and cloned into pET24b(+) or pET28b(+) vectors. The recombinant *C*-His$_6$-LmbC, *N*-His$_6$-LmbE, *N*-His$_6$-LmbF and *N*-His$_6$-LmbG were isolated as soluble proteins (Figure 5.5), but LmbD, LmbT and LmbV were expressed only as inclusion bodies. The genes *ccbD*, *ccbF*, *ccbT* and *ccbV* encoding the homologous proteins in celesticetin biosynthetic pathway were then amplified and cloned into pET28b(+). The *N*-His$_6$-CcbD, *N*-His$_6$-CcbF, *N*-His$_6$-CcbT and *N*-His$_6$-CcbV were obtained in soluble forms (Figure 5.5).

5.3.2 Sulfur Transfer

To examine the hypothesis that LmbT and/or LmbV are responsible for the sulfur transfer, the chemically prepared GDP-aminooctose **5-19** was then incubated with the purified CcbT and CcbV and the potential sulfur sources including L-cysteine (**5-2**), GSH (**5-15**) and MSH (**5-16**). The HPLC analysis of the reaction mixtures was shown in Figure 5.6. Neither the consumption of the substrate **5-19** nor the formation of GDP was observed in all samples subjected to different testing conditions. The negative results suggested that the attachment of the amino acid moiety may precede the sulfur transfer reaction in the proposed pathway (Figure 5.4). However, it is also conceivable that the LMW thiols tested in the experiments are not relevant sulfur donors in lincomycin A biosynthesis.

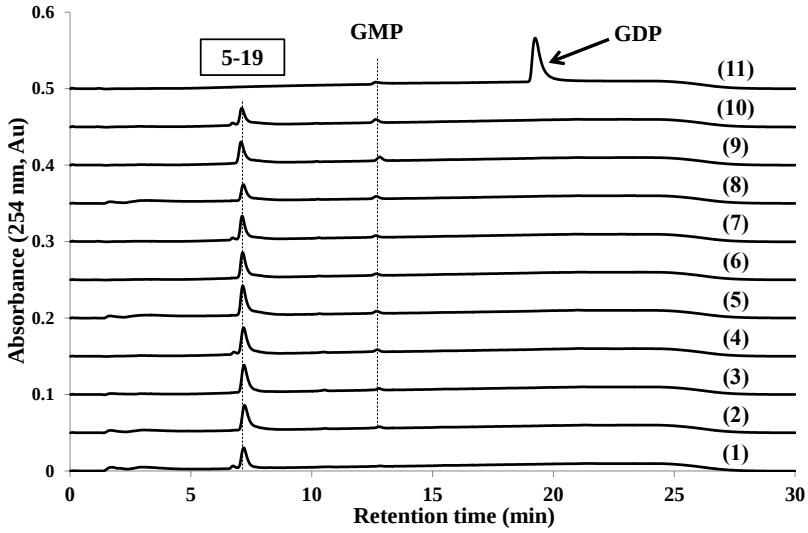


Figure 5.6: HPLC analysis of CcbT and CcbV reactions using **5-19** and different thiols as substrates. All reaction mixtures contained DTT and **5-19**. (1) Control reaction with L-cysteine, GSH and MSH. (2) CcbT-reaction wih MSH, (3) L-cysteine, (4) GSH. (5) CcbV-reaction wih MSH, (6) L-cysteine, (7) GSH. (8) CcbT and CcbV-reaction wih MSH, (9) L-cysteine, (10) GSH. (11) GDP standard.

5.3.3 Amide Bond Formation

Characterization of the putative T-domain subunits in lincosamide synthetase

In Section 2.3.1, the resemblance of the amino acid incorporation in lincomycin A formation and that of nonribosomal peptide (NRP) biosynthesis was discussed. According to the results of sequence analysis, the previously missing T- and C-domain are predicted to be encoded by *lmbN* and *lmbD*, respectively. The T-domain of LmbN and the C-domain of LmbD inconjunction with LmbC, which has recetnly been identified as the A-domain,[137] are expected to form a protein complex, lincosamide synthetase. Together, they catalyze the amide bond formation between the propylproline (**5-33**) and the octose moiety (**5-36**) (see Figure 5.8).

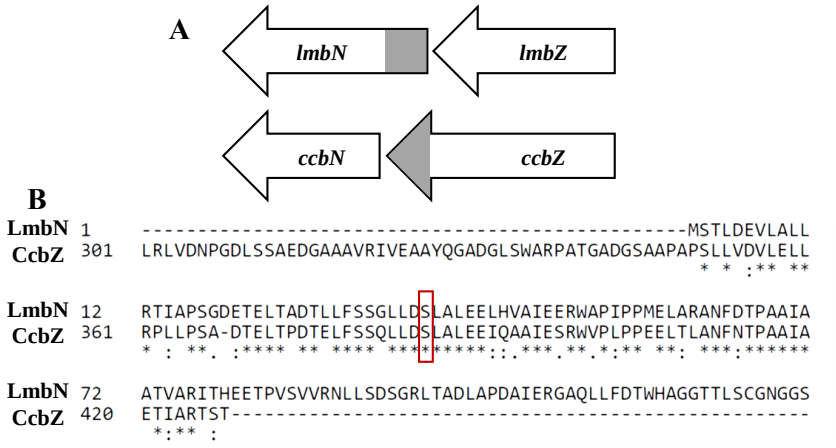


Figure 5.7: Alignment of the gene and amino acid sequences of the putative T-domains. (A) Alignment of genes *lmbN-lmbZ* and *ccbN-ccbZ*. The regions encoding the putative T-domains are shown in gray. (B) The overlapping region of the amino acids sequences of LmbN and CcbZ. The serine residue conserved in all acyl- and peptidyl carrier proteins is highlighted with red box.

Although the isomerase activity of LmbN has been established in Section 2.3.3 (page 72), a closer comparison of the amino acid sequence revealed that a putative T-domain is enbeded at the *C*-terminus of LmbN (encoded by the 5′-terminus of *lmbN*, Figure 5.7). Interestingly, the putative T-domain in celesticetin biosynthetic gene cluster is found to be encoded at the 3′-terminus of *ccbZ*, which is the down-

stream gene of *ccbN* (the counterpart of *lmbN*). The oxidoreductase activity of CcbZ has also been characterized in Section 4.3.4 (page 128). The alignment of LmbN and CcbZ revealed a overlapping region of 79 amino acid residues. A conserved serine residue can be identified when aligning this fragment with known acyl- or peptidyl-carrier proteins in database. Therefore, this fragment may function as the T-domain subunit of lincosamide synthetase.

The chemistry of T-domain in lincosamide synthetase is expected to be analogous to that of T-domain in nonribosomal peptide synthetase (NRPS) assembly line. That is, the apo-T-domain (**5-30**) must be first primed by post-translational modification of a conserved serine residue with 4′-phosphopantetheine group (Figure 5.8). The nucleophilic thiolate of **5-32** then attacks an adenylated propylproline (**5-34**), selectively activated by LmbC, to form a reactive thioester intermediate (**5-35**). This protein-bound amino acid **5-35** is subsequently utilized by LmbD to form an amide linkage with the aminosugar (**5-36**).

Therefore, the purified *C*-His_6-LmbN and *N*-His_6-CcbZ were analyzed by LC-ESI-MS for molecular weight determination (Figure 5.9). The major form of as-purified LmbN showed molecular weight of 31234 Da, ∼125 Da less than the calculated number 31359 Da. It is likely that the *C*-termial His-tag of the protein was nibbled during the expression in *E. coli*, so the predominate form of LmbN actually has less than six histidine residues. The other species (observed mass: 31573 Da) of the as-purified LmbN corresponds in molecular mass to a 4′-phosphopantetheine modified protein (with an extra mass of 339 Da). This observation indicated that part of the as-purified LmbN was already modified in the *E. coli* host and existed in holo-form. The as-purified CcbZ also contained two species with mass values of 48188 and 48366 Da, respectively (calcd. 47940 Da). The extra masses were most likely attributed to the α-*N*-phosphogluconoylation and -gluconoylation of the *N*-terminal His_6 tags, respectively.[182] Unlike LmbN, the predominant form of

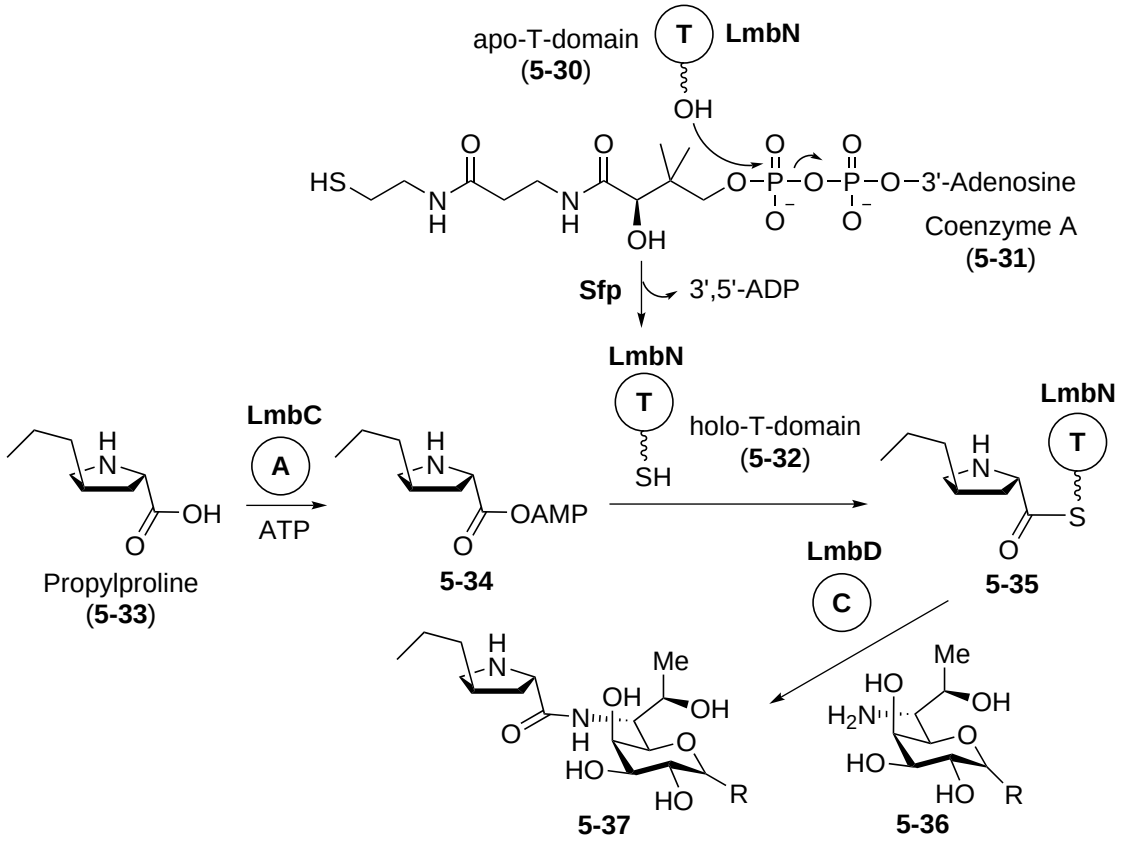


Figure 5.8: Proposed mechanism of the amide bond formation catalyzed by lincosamide synthetase. Post-translational modification of apo-T-domain (**5-30**) by phosphopantetheinyl transferase, Sfp, to produce holo-T-domain (**5-32**). The A-domain, LmbC, catalyzes the adenylation of propylproline (**5-33**) and the resulting adenylated intermediate (**5-34**) then reacts with **5-32** to afford the acyl-LmbN adduct (**5-35**). The C-domain, LmbD, catalyzes the C–N bond formation between the electrophilic peptidyl-protein (**5-35**) and the nucleophilic aminosugar (**5-36**).

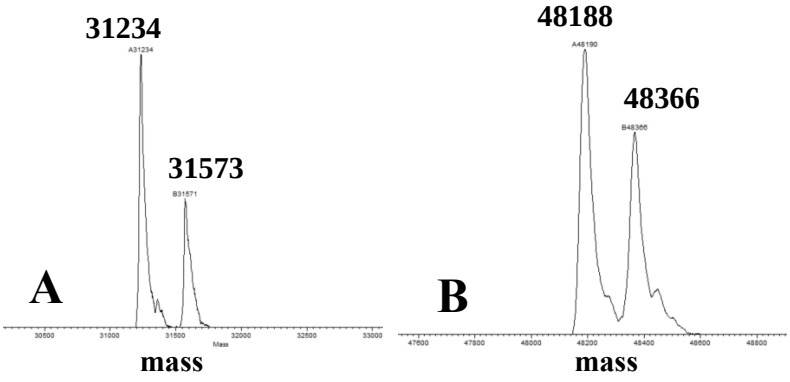


Figure 5.9: Deconvoluted ESI-MS analysis of the (A) *C*-His$_6$-LmbN (calcd. 31359 Da; obsd. 31234 and 31573 Da) and (B) *N*-His$_6$-CcbZ (calcd. 47940 Da; obsd. 48188 and 48366 Da).

the as-purified CcbZ was apo-T-domain.

Functional characterization of lincosamide synthetase

To verify the proposed function of lincosamide synthetase, the potential substrates, GDP-aminoocotse (**5-19**) and methylthiolincosamide (MTL, **5-23**), were incubated with LmbC, CcbD (functional identical to LmbD), LmbN, ATP and chemically synthesized propylproline (**5-33**). The chemical synthesis of propylproline was performed by another graduate student, Richiro Ushimaru, using a known procedure.[183] The reaction mixtures were added coenzyme A and Sfp (phosphopantetheinyl transferase) to ensure the phosphopantetheinylation of the carrier protein, LmbN.[184] When using MTL (**5-23**) as the substrate, the production of the amide bond-containing molecule **5-28** was detected by LC-MS (Figure 5.10). However, the installation of the amino acid derivative was not observed when using GDP-aminosugar (**5-19**) as the substrate (data not shown).

Since LmbC has been shown to also catalyze adenylation using L-proline with $\sim 10^3$ times lower affinity than propylproline (**5-34**),[137] the reaction using L-proline as the amino acid donor was carried out to examine whether the T- and C-domains can accept the adenylated L-proline. Based on the LC-MS analysis of the enzymatic reaction, formation of proline-sugar conjugate **5-38** was observed (Figure 5.11). Fusaric acid (**5-39**) and picolinic acid (**5-40**), structurally related to propylproline (Figure 5.12), were also tested for their activities toward lincosamide synthetase. The LC-MS analysis showed no incorporation was observed in both cases (data not shown). The relatively strict substrate selectivity of LmbC and the relaxed substrate specificity of LmbN and CcbD are consistent with the general features of NRPS. The catalytic A-domain usually serves as the gate-keeper that prevent mis-incorporation of an incorrect amino acid, while the T- and C-domains have broader substrate specificity and accept noncognate donors.[136]

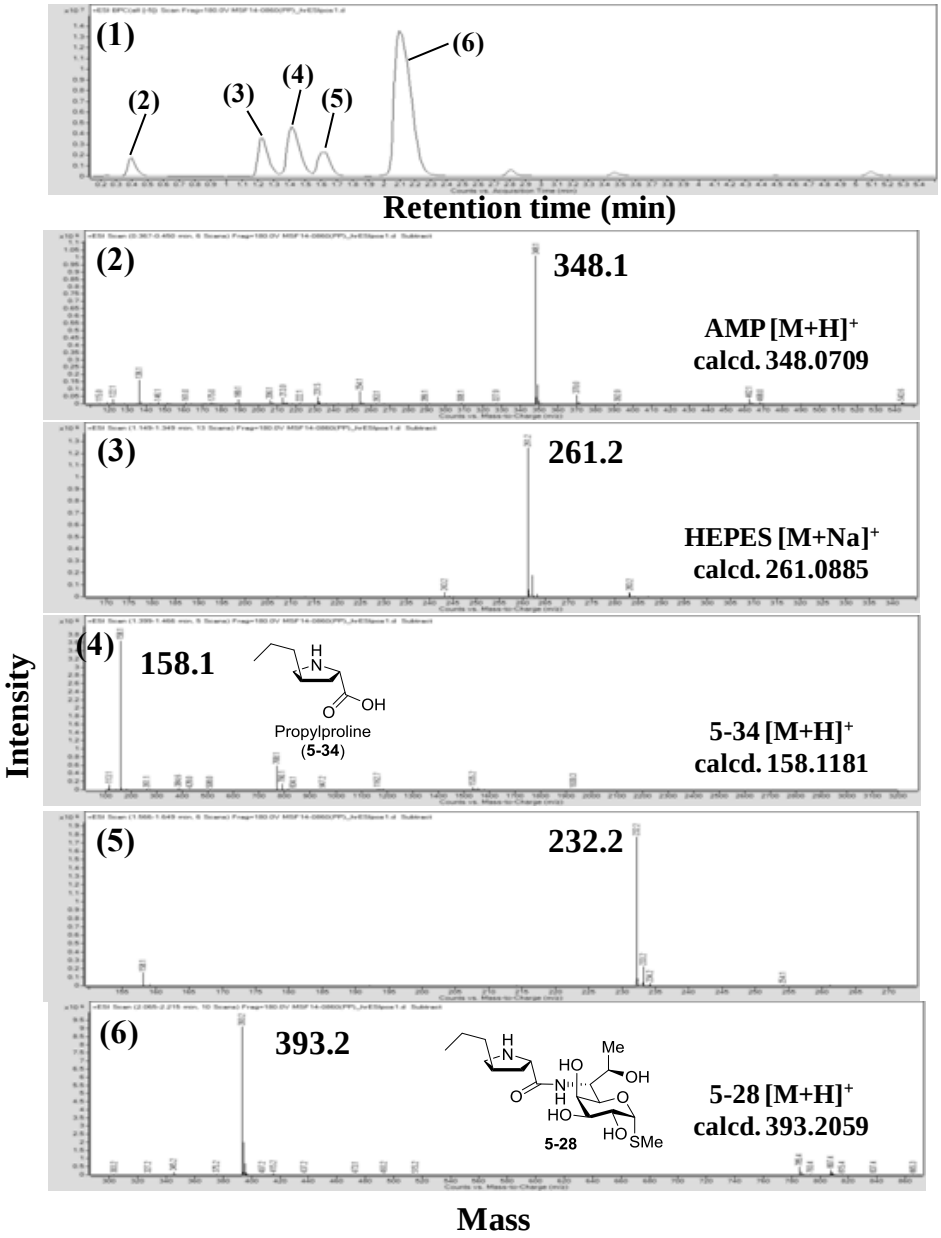


Figure 5.10: LC-MS analysis of the propylproline incorporation into MTL. (1) Full-mass chromatogram. (2) ESI-MS spectrum of species eluted at 0.37–0.45 min, (3) 1.15–1.35 min, (4) 1.40–1.47 min, (5) 1.57–1.65 min, and (6) 2.07–2.22 min.

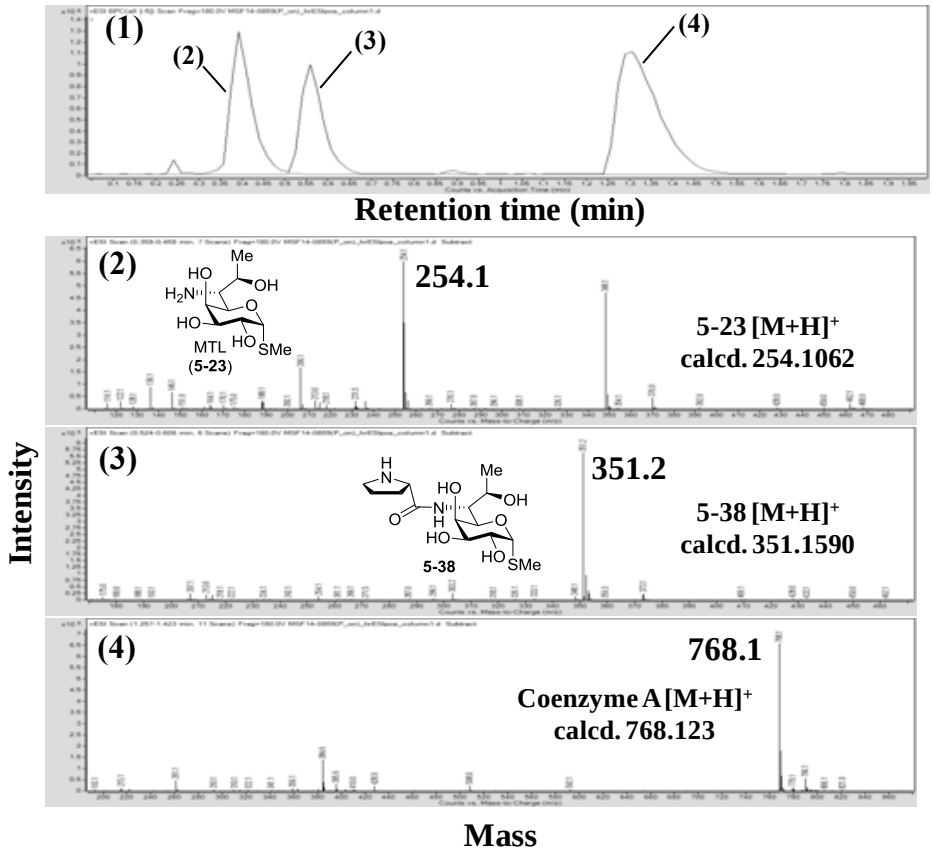


Figure 5.11: LC-MS analysis of the L-proline incorporation into MTL. (1) Full-mass chromatogram. ESI-MS spectrum of species eluted at (2) 0.36–0.46 min, (3) 0.52–0.61 min, and (4) 1.26–1.42 min.

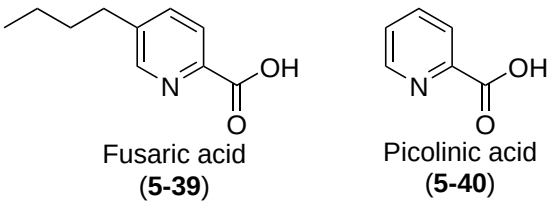


Figure 5.12: Fusaric acid (**5-39**) and picolinic acid (**5-40**).

Although the T-domain of typical NRPS are often fused to functionally related enzymes, such as A- or C-domains, this possibility is not the case in lincomycin A biosynthesis. The fusion partners of the two T-domains are either an isomerase (LmbN) or an oxidoreductase (CcbZ), which are involved in the construction and modification of the aminooctose. Since both fused proteins were shown to be active in the amide bond formation, we decided to investigate whether the thiolation activity is dependent on the fusion partners. Accordingly, the standalone T-domain encoded in *lmbN* and *ccbZ* was individually cloned, overexpressed and purified in *N*-His-tagged form (Figure 5.13). Both *N*-His_6-LmbN-T and *N*-His_6-CcbZ-T were capable of catalyzing the formation of **5-28** using propylproline (**5-38**) and MTL (**5-23**). These results indicated that the thiolation and isomerization/oxidoreduction are independent of each other. Clearly, LmbN and CcbZ serve as bi-functional enzymes, both the fusion partners and thiolation domains of which are essential for the lincomycin A or celesticetin biosynthesis. Although fusion of the T-domain to enzymes performing indirect functions is not common, few examples related to nonribosomal peptide biosynthesis have been reported.[185,186]

Because propylproline (**5-38**) could be attached to methylthiolincosamide (**5-23**) but not GDP-aminooctose **5-19**, it is thus confirmed that the sulfur incorporation or the detachment of GDP group must proceed prior to the amide bond formation. The failure to observe sulfur transfer using GDP-aminosugar as substrate may due to the improper reaction conditions or the incorrect sulfur donors.

5.3.4 Recently Reported Sulfur Transfer in Lincomycin A Biosynthesis

Very recently, Zhao *et al.* have successfully demonstrated the enzymatic sulfur incorporation in lincomycin A biosynthesis (Figure 5.15).[187] The authors identified an unprecedented coupling of two bacterial thiols, mycothiol (**5-16**) and ergothioneine (EGT, **5-18**), to install the sulfur atom into the lincosamide octose.

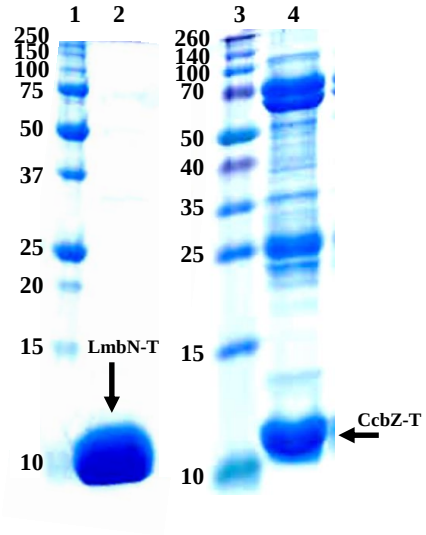


Figure 5.13: SDS-PAGE gel of *N*-His_6-LmbN-T (10.6 kDa, lane 2) and *N*-His_6-CcbZ-T (10.6 kDa, lane 4). The molecular markers are shown in lanes 1 and 3. Minor impurities were possibly resulted from the non-specific interactions with Ni-NTA resins.

The GDP-activated aminooctose **5-19** is first linked with EGT (**5-18**) via the glycosylation catalyzed by LmbT. The EGT-octose conjugate (**5-41**) then serves as the substrate for the condensation with the propylproline moiety through the catalysis of lincosamide synthetase (LmbC, CcbD and LmbN). Mycothiol (**5-16**), in turn, activating by CcbV, participate in the second *S*-glycosylation to exchange with EGT to yield the MSH-octose intermediate (**5-43**), which is subsequently hydrolyzed by LmbE.

This report substantiated the hypothesis that mycothiol (**5-16**) is the ultimate sulfur source for lincomcyin A biosynthesis. Moreover, the discovery of the EGT-mediated thiol exchange explains the observation that the displacement of GDP did not occur when GDP-aminooctose (**5-19**) was incubated with MSH (**5-16**) and the glycosyltransferase CcbT. Although our data suggested that the lincosamide synthetase can utilize methylthiolincosamide (**5-23**) for the amide bond formation, it is probably resulted from the relaxed substrate specificity of LmbN and CcbD.

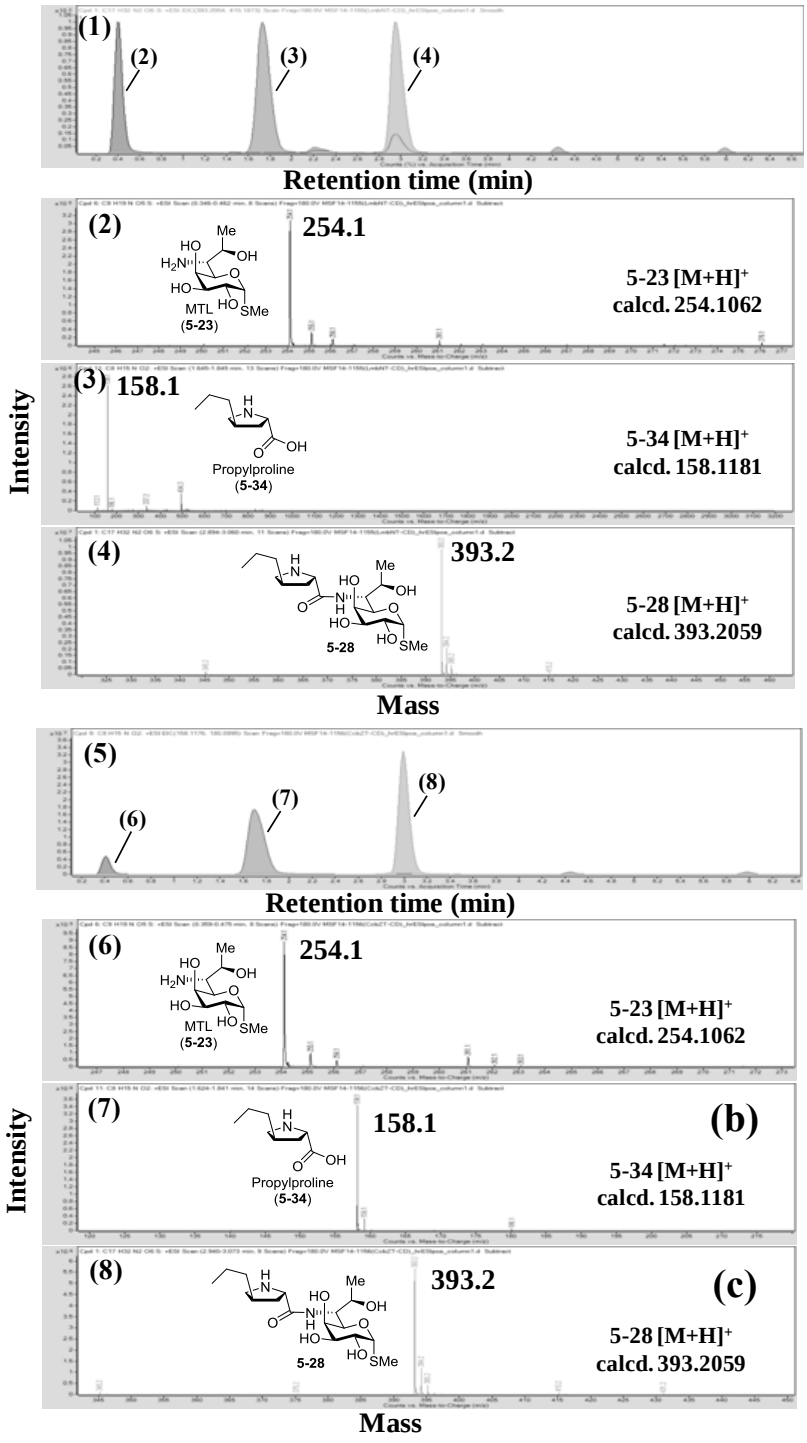


Figure 5.14: LC-MS analysis of the propylproline incorporation into MTL catalyzed by *N*-His$_6$-LmbN-T (1–4) and *N*-His$_6$-CcbZ-T (5–8). (1) Full-mass chromatogram of LmbN-T-reaction mixture. ESI-MS spectrum of species eluted at (2) 0.35–0.46 min, (3) 1.65–1.85 min, and (4) 2.89–3.06 min. (1) Full-mass chromatogram of CcbZ-T-reaction mixture. ESI-MS spectrum of species eluted at (2) 0.36–0.48 min, (3) 1.62–1.84 min, and (4) 2.94–3.07 min.

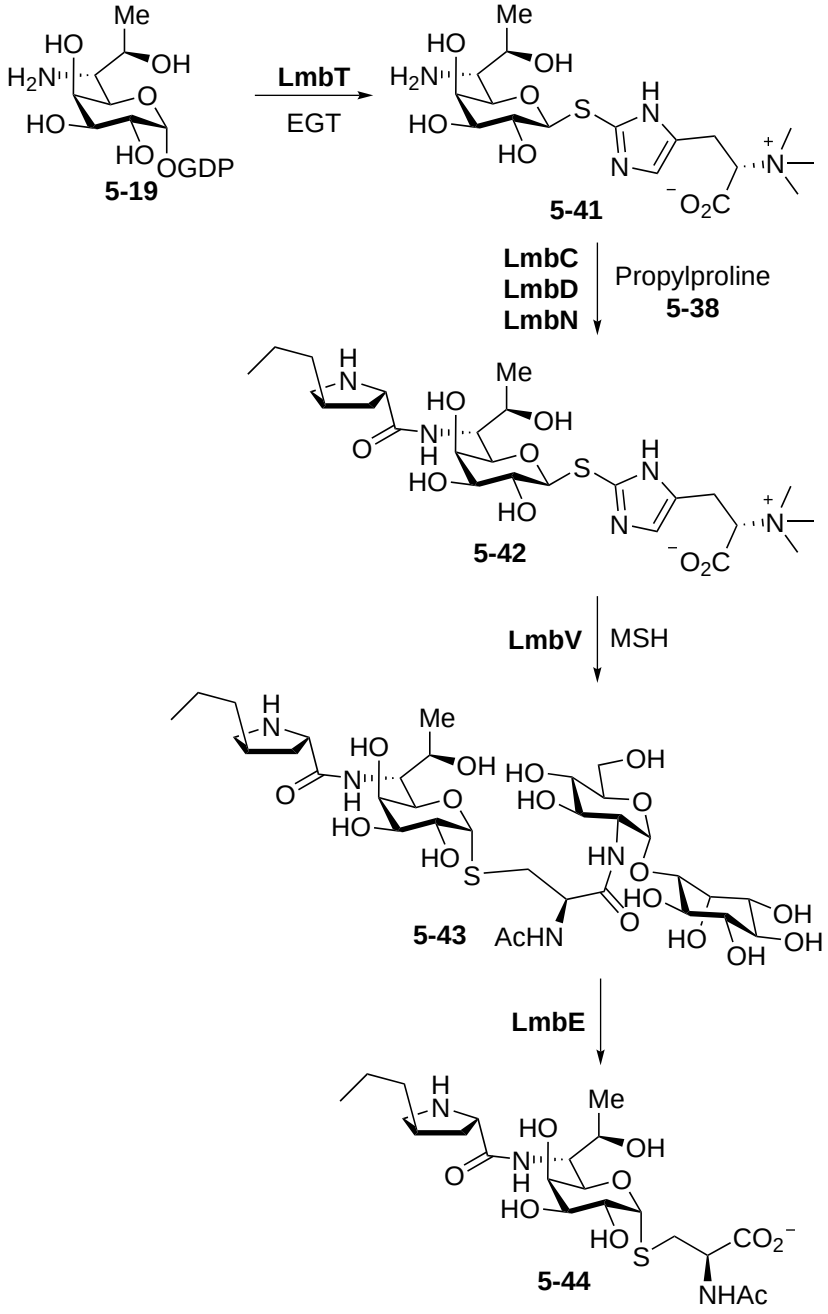


Figure 5.15: The participation of EGT and MSH in the sulfur incorporation of the lincomycin A biosynthetic pathway.

5.4 Conclusion

In this chapter, the sulfur transfer and amide bond formation in lincomycin A biosynthetic pathway was investigates. Although our attempts to demonstrate the mycothiol incorporation into GDP-aminosugar were unsuccessful, a recent report about the sequential thiol exchange between ergothioneine and mycothiol as the mechanism of sulfur incorporation in lincomycin A biosynthesis complemented our results and validated our hypothesis. Since ionic sulfur transfer involving low-molecular weight thiols is rare, this finding expands our knowledge regarding the biosynthesis of sulfur-containing natural products.

Moreover, the identification and characterization of a multi-component protein comples, lincosamide synthetase, reveals a unique system catalyzing the condensation of an unusual amino acid, propylproline, and an aminosugar, methylthiolincosamide. Despite the fairly low homology of LmbD/CcbD to any sequences documented in database, fragment-based approach combined with the knowledge of enzyme mechanism allows us to predict the catalytic function of LmbD/CcbD. The discovery of the hidden thiolation domains of the *N*-terminus of LmbN and C-terminus of CcbZ demonstrates the mobility of the carrier protein in a NRPS-like biosynthetic machinery. This finding holds a great potential for metabolic engineering of the lincomycin A biosynthesis. The relatively flexible substrate specificity of the T- and C-domains of lincosamide synthetase may allow replacing the proline derivatives to other moieties. It may be easily accomplished by swapping the proline specific A-domain (LmbC) to other carboxylic acid adenylating enzymes.

Overall it is the novelty of chemical structure and the complexity of the proposed biosynthetic pathway of lincomycin A prompted us to investigate the construction of its unique thiosugar component. For fifty years since the first isolation

of lincomycin A, the accumulated information from the feeding experiments with radio-labeled compounds, the discovery of the biosynthetic gene cluster, and the results of gene disruption studies have paved the road for us to reconstitute the lincomycin A biosynthetic pathway *in vitro* via biochemical characterization of each enzyme using chemically synthesized putative substrate intermediates. Rooting in two primary metabolites, D-sedoheptulose 7-phsophate and D-ribose 5-phosphate, from the pentose phosphate pathway, the biosynthetic pathway of lincomycin A consolidates the C-1 activation strategy observed in the construction of bacterial cell wall and the GDP-octose fabrication evolved from the NDP-hexose modification. The unprecedent sulfur transfer involving the sequential substitution of two thiols along with the distinct NRPS system catalyzed coupling of propylproline to the thiosugar unit represents an elegantly programmed biosynthetic machinery emerging from evolution. The biosynthetic pathway of a lincomycin A analog, celesticetin, was also proposed in this study. The chemical synthesis of the key intermediates and the characterization of the subsequent enzymatic reactions are currently in progress to elucidate the diverting point in the biosynthetic pathways of these two thiosugar-containing natural products.

Bibliography

[1] Elshahawi, S. I.; Shaaban, K. A.; Kharel, M. K.; Thorson, J. S. A comprehensive review of glycosylated bacterial natural products. *Chem. Soc. Rev.* **2015**, *ASAP*.

[2] Chen, X. Fermenting Next Generation Glycosylated Therapeutics. *ACS Chem. Biol.* **2011**, *6*, 14–17.

[3] Timmons, S. C.; Thorson, J. S. Increasing carbohydrate diversity via amine oxidation: aminosugar, hydroxyaminosugar, nitrososugar, and nitrosugar biosynthesis in bacteria. *Curr. Opin. Chem. Biol.* **2008**, *12*, 297–305.

[4] Nedal, A.; Zotchev, S. B. Biosynthesis of deoxyaminosugars in antibiotic-producing bacteria. *Appl. Microbiol. Biotechnol.* **2004**, *64*, 7–15.

[5] Parry, R. J. In *Comprehensive Natural Products Chemistry*; Meth-Cohn, O., Barton, S. D., Nakanishi, K., Eds.; Pergamon: Oxford, 1999; pp 825–863.

[6] Witczak, Z.; Culhane, J. Thiosugars: new perspectives regarding availability and potential biochemical and medicinal applications. *Appl. Microbiol. Biotechnol.* **2005**, *69*, 237–244.

[7] Neumann, C. S.; Fujimori, D. G.; Walsh, C. T. Halogenation Strategies In Natural Product Biosynthesis. *Chem. Biol.* **2008**, *15*, 99–109.

[8] Feldmann, J.; Krupp, E. M. Critical review or scientific opinion paper: Arsenosugars–a class of benign arsenic species or justification for developing partly speciated arsenic fractionation in foodstuffs? *Anal. Bioanal. Chem.* **2011**, *399*, 1735–1741.

[9] Metcalf, W. W.; van der Donk, W. A. Biosynthesis of Phosphonic and Phosphinic Acid Natural Products. *Annu. Rev. Biochem.* **2009**, *78*, 65–94.

[10] Thibodeaux, C. J.; Melancon, C. E.; Liu, H.-w. Unusual sugar biosynthesis and natural product glycodiversification. *Nature* **2007**, *446*, 1008–1016.

[11] Thibodeaux, C.; Melanon, C.; Liu, H.-w. Natural-Product Sugar Biosynthesis and Enzymatic Glycodiversification. *Angew. Chem. Int. Ed.* **2008**, *47*, 9814–9859.

[12] Yuasa, H.; Izumi, M.; Hashimoto, H. Thiasugars as Potential Glycosidase Inhibitor. *ChemInform* **2002**, *33*, 250–250.

[13] Etoh, H.; Iguchi, M.; Nagasawa, T.; Tani, Y.; Yamada, H.; Fukami, H. Structures of rhodonocardins produced by a *Nocardia* sp. *Agric. Biol. Chem.* **1987**, *51*, 1819–1824.

[14] Capon, R. J.; MacLeod, J. K. 5-Thio-D-mannose from the marine sponge *Clathria pyramida* (Lendenfeld). The first example of a naturally occurring 5-thiosugar. *J. Chem. Soc., Chem. Commun.* **1987**, 1200–1201.

[15] Benz, G.; Schröder, T.; Kurz, J.; Wünsche, C.; Karl, W.; Steffens, G.; Pfitzner, J.; Schmidt, D. Konstitution der Desferriform der Albomycine δ1, δ2, ϵ. *Angew. Chem. Int. Ed.* **1982**, *21*, 1322–1335.

[16] Benz, G.; Born, L.; Brieden, M.; Grosser, R.; Kurz, J.; Paulsen, H.; Sinnwell, V.; Webber, B. Albomycine, II. Absolute Konfiguration der Desferriform der Albomycine. *Liebigs Ann. Chem.* **1984**, *1984*, 1408–1423.

[17] Braun, V.; Pramanik, A.; Gwinner, T.; Kberle, M.; Bohn, E. Sideromycins: tools and antibiotics. *BioMetals* **2009**, *22*, 3–13.

[18] Stefanska, A. L.; Fulston, M.; Houge-Frydrych, C. S.; Jones, J. J.; Warr, S. R. A Potent Seryl tRNA Synthetase Inhibitor SB-217452 Isolated from a *Streptomyces species*. *J. Antibiot.* **2000**, *53*, 1346–1353.

[19] Paulsen, H.; Brieden, M.; Benz, G. Verzweigte und kettenverlngerte Zucker, XXXI. Synthese des Sauerstoffanalogons der Desferriform von 1-Albomycin. *Liebigs Ann. Chem.* **1987**, *1987*, 565–575.

[20] Zeng, Y.; Kulkarni, A.; Yang, Z.; Patil, P. B.; Zhou, W.; Chi, X.; Van Lanen, S.; Chen, S. Biosynthesis of Albomycin δ_2 Provides a Template for Assembling Siderophore and Aminoacyl-tRNA Synthetase Inhibitor Conjugates. *ACS Chem. Biol.* **2012**, *7*, 1565–1575.

[21] Jarrett, J. T. The novel structure and chemistry of iron-sulfur clusters in the adenosylmethionine-dependent radical enzyme biotin synthase. *Arch. Biochem. Biophys.* **2005**, *433*, 312–321.

[22] Booker, S. J.; Cicchillo, R. M.; Grove, T. L. Self-sacrifice in radical *S*-adenosylmethionine proteins. *Curr. Opin. Chem. Biol.* **2007**, *11*, 543–552.

[23] Frey, P. A.; Hegeman, A. D.; Ruzicka, F. J. The Radical SAM Superfamily. *Crit. Rev. Biochem. Mol. Biol.* **2008**, *43*, 63–88.

[24] Okabe, T.; Suda, H.; Sato, F.; Okanishi, M. JP Patent JP 02-16894 A, 1990.

[25] Sasaki, E.; Liu, H.-w. Mechanistic Studies of the Biosynthesis of 2-Thiosugar: Evidence for the Formation of an Enzyme-Bound 2-Ketohexose Intermediate in BexX-Catalyzed Reaction. *J. Am. Chem. Soc.* **2010**, *132*, 15544–15546.

[26] Begley, T. P.; Ealick, S. E.; McLafferty, F. W. Thiamine biosynthesis: still yielding fascinating biological chemistry. *Biochem. Soc. Trans.* **2012**, *40*, 555–560.

[27] Sasaki, E.; Zhang, X.; Sun, H. G.; Lu, M.-Y. J.; Liu, T.-l.; Ou, A.; Li, J.-y.; Chen, Y.-h.; Ealick, S. E.; Liu, H.-w. Co-opting sulphur-carrier proteins from primary metabolic pathways for 2-thiosugar biosynthesis. *Nature* **2014**, *520*, 427–431.

[28] Mason, D. J.; Dietz, A.; DeBoer, C. Lincomycin, a new antibiotic. I. Discovery and biological properties. *Antimicrob. Agents Chemother.* **1962**, 554–559.

[29] Mason, D. J.; Lewis, C. Biological activity of the lincomycin related antibiotics. *Antimicrob. Agents Chemother.* **1964**, 7–12.

[30] Dunkle, J. A.; Xiong, L.; Mankin, A. S.; Cate, J. H. D. Structures of the *Escherichia coli* ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 17152–17157.

[31] Hoeksema, H.; Bannister, B.; Birkenmeyer, R. D.; Kagan, F.; Magerlein, B. J.; MacKellar, F. A.; Schroeder, W.; Slomp, G.; Herr, R. R. Chemical Studies on Lincomycin. I. The Structure of Lincomycin. *J. Am. Chem. Soc.* **1964**, *86*, 4223–4224.

[32] Hanada, M.; Tsunakawa, M.; Tomita, K.; Tsukiura, H.; Kawaguchi, H. Antibiotic Bu-2545, a new member of the celesticetin-lincomycin class. *J. Antibiot.* **1980**, *33*, 751–753.

[33] Hoeksema, H.; Crum, G. F.; DeVries, W. H. Isolation and purification of celesticetin. *Antibiot. Ann.* **1955**, *2*, 837–841.

[34] Lee, M. D.; Dunne, T. S.; Chang, C. C.; Ellestad, G. A.; Siegel, M. M.; Morton, G. O.; McGahren, W. J.; Borders, D. B. Calichemicins, a novel family of antitumor antibiotics. 2. Chemistry and structure of calichemicin γ_1^I. *J. Am. Chem. Soc.* **1987**, *109*, 3466–3468.

[35] McDonald, L. A.; Capson, T. L.; Krishnamurthy, G.; Ding, W.-D.; Ellestad, G. A.; Bernan, V. S.; Maiese, W. M.; Lassota, P.; Kramer, R. A.; Ireland, C. M. Namenamicin, a New Enediyne Antitumor Antibiotic from the Marine Ascidian Polysyncraton lithostrotum. *J. Am. Chem. Soc.* **1996**, *118*, 10898–10899.

[36] Golik, J.; Clardy, J.; Dubay, G.; Groenewold, G.; Kawaguchi, H.; Konishi, M.; Krishnan, B.; Ohkuma, H.; Saitoh, K.; Doyle, T. W. Esperamicins, a novel class of potent antitumor antibiotics. 2. Structure of esperamicin X. *J. Am. Chem. Soc.* **1987**, *109*, 3461–3462.

[37] Oku, N.; Matsunaga, S.; Fusetani, N. Shishijimicins A–C, Novel Enediyne Antitumor Antibiotics from the Ascidian *Didemnum proliferum*. *J. Am. Chem. Soc.* **2003**, *125*, 2044–2045.

[38] Ahlert, J.; Shepard, E.; Lomovskaya, N.; Zazopoulos, E.; Staffa, A.; Bachmann, B. O.; Huang, K.; Fonstein, L.; Czisny, A.; Whitwam, R. E.; Farnet, C. M.; Thorson, J. S. The Calicheamicin Gene Cluster and Its Iterative Type I Enediyne PKS. *Science* **2002**, *297*, 1173–1176.

[39] Agerbirk, N.; Olsen, C. E. Glucosinolate structures in evolution. *Phytochemistry* **2012**, *77*, 16–45.

[40] Fahey, J. W.; Zalcmann, A. T.; Talalay, P. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* **2001**, *56*, 5–51.

[41] Halkier, B. A.; Gershenzon, J. Biology and Biochemistry of Glucosinolates. *Annu. Rev. Plant Biol.* **2006**, *57*, 303–333.

[42] Sonderby, I. E.; Geu-Flores, F.; Halkier, B. A. Biosynthesis of glucosinolates - gene discovery and beyond. *Trends Plant Sci.* **2010**, *15*, 283–290.

[43] Schlaeppi, K.; Bodenhausen, N.; Buchala, A.; Mauch, F.; Reymond, P. The glutathione-deficient mutant *pad2-1* accumulates lower amounts of glucosinolates and is more susceptible to the insect herbivore *Spodoptera littoralis*. *Plant J.* **2008**, *55*, 774–786.

[44] Bednarek, P.; Pislewska-Bednarek, M.; Svatos, A.; Schneider, B.; Doubsky, J.; Mansurova, M.; Humphry, M.; Consonni, C.; Panstruga, R.; Sanchez-Vallet, A.; Molina, A.; Schulze-Lefert, P. A Glucosinolate Metabolism Pathway in Living Plant Cells Mediates Broad-Spectrum Antifungal Defense. *Science* **2009**, *323*, 101–106.

[45] Geu-Flores, F.; Nielsen, M. T.; Nafisi, M.; Moldrup, M. E.; Olsen, C. E.; Motawia, M. S.; Halkier, B. A. Glucosinolate engineering identifies a γ-glutamyl peptidase. *Nat. Chem. Biol.* **2009**, *5*, 575–577.

[46] Yoshikawa, M.; Murakami, T.; Shimada, H.; Matsuda, H.; Yamahara, J.; Tanabe, G.; Muraoka, O. Salacinol, potent antidiabetic principle with

unique thiosugar sulfonium sulfate structure from the Ayurvedic traditional medicine *Salacia reticulata* in Sri Lanka and India. *Tetrahedron Lett.* **1997**, *38*, 8367–8370.

[47] Yoshikawa, M.; Murakami, T.; Yashiro, K.; Matsuda, H. Kotalanol, a Potent α-Glucosidase Inhibitor with Thiosugar Sulfonium Sulfate Structure, from Antidiabetic Ayurvedic Medicine *Salacia reticulata*. *Chem. Pharm. Bull.* **1998**, *46*, 1339–1340.

[48] Yoshikawa, M.; Xu, F.; Nakamura, S.; Wang, T.; Matsuda, H.; Tanabe, G.; Muraoka, O. Salaprionol and Ponkoranol with Thiosugar Sulfonium Sulfate Structure from *Salacia prinoides* and α-Glucosidase Inhibitory Activity of Ponkoranol and Kotalanol Desulfate. *Heterocycles* **2008**, *75*, 1397–1405.

[49] Mohan, S.; Pinto, B. M. Zwitterionic glycosidase inhibitors: salacinol and related analogues. *Carbohydr. Res.* **2007**, *342*, 1551–1580.

[50] Benson, A. A.; Daniel, H.; Wiser, R. A Sulfolipid in Plants. *Proc. Natl. Acad. Sci. U.S.A.* **1959**, *45*, 1582–1587.

[51] Block, M. A.; Dorne, A. J.; Joyard, J.; Douce, R. Preparation and characterization of membrane fractions enriched in outer and inner envelope membranes from spinach chloroplasts. II. Biochemical characterization. *J. Biol. Chem.* **1983**, *258*, 13281–13286.

[52] Shibuya, I.; Yagi, T.; Benson, A. A. *Japanese Society of Plant Physiology*; University of Tokyo Press: Tokyo, 1963; pp 627–636.

[53] Benning, C.; Somerville, C. R. Identification of an operon involved in sulfolipid biosynthesis in *Rhodobacter sphaeroides*. *J. Bacteriol.* **1992**, *174*, 6479–6487.

[54] Pugh, C. E.; Roy, A. B.; Hawkes, T.; Harwood, J. L. A new pathway for the synthesis of the plant sulpholipid, sulphoquinovosyldiacylglycerol. *Biochem. J.* **1995**, *309*, 513–519.

[55] Sanda, S.; Leustek, T.; Theisen, M. J.; Garavito, R. M.; Benning, C. Recombinant Arabidopsis SQD1 Converts UDP-glucose and Sulfite to the Sulfolipid Head Group Precursor UDP-sulfoquinovose in Vitro. *J. Biol. Chem.* **2001**, *276*, 3941–3946.

[56] Raetz, C. R. H.; Whitfield, C. Lipopolysaccharide Endotoxins. *Annu. Rev. Biochem.* **2002**, *71*, 635–700.

[57] Kosma, P. Occurrence, Synthesis and Biosynthesis of Bacterial Heptoses. *Curr. Org. Chem.* **2008**, *12*, 1021–1039.

[58] Messner, P.; Steiner, K.; Zarschler, K.; Schaffer, C. S-layer nanoglycobiology of bacteria. *Carbohydr. Res.* **2008**, *343*, 1934–1951.

[59] Niu, G.; Tan, H. Nucleoside antibiotics: biosynthesis, regulation, and biotechnology. *Trends Microbiol.* **2015**, *23*, 110–119.

[60] Rodolis, M. T.; Mihalyi, A.; Ducho, C.; Eitel, K.; Gust, B.; Goss, R. J. M.; Bugg, T. D. H. Mechanism of action of the uridyl peptide antibiotics: an unexpected link to a protein-protein interaction site in translocase MraY. *Chem. Commun.* **2014**, *50*, 13023–13025.

[61] Funabashi, M.; Baba, S.; Nonaka, K.; Hosobuchi, M.; Fujita, Y.; Shibata, T.; Van Lanen, S. G. The biosynthesis of liposidomycin-like A-90289 antibiotics featuring a new type of sulfotransferase. *ChemBioChem* **2010**, *11*, 184–190.

[62] Kaysser, L.; Lutsch, L.; Siebenberg, S.; Wemakor, E.; Kammerer, B.; Gust, B. Identification and manipulation of the caprazamycin gene cluster lead to new simplified liponucleoside antibiotics and give insights into the biosynthetic pathway. *J. Biol. Chem.* **2009**, *284*, 14987–14996.

[63] Kaysser, L.; Siebenberg, S.; Kammerer, B.; Gust, B. Analysis of the liposidomycin gene cluster leads to the identification of new caprazamycin derivatives. *ChemBioChem* **2010**, *11*, 191–196.

[64] Cheng, L.; Chen, W.; Zhai, L.; Xu, D.; Huang, T.; Lin, S.; Zhou, X.; Deng, Z. Identification of the gene cluster involved in muraymycin biosynthesis from *Streptomyces* sp. NRRL 30471. *Mol. BioSyst.* **2011**, *7*, 920–927.

[65] Yang, Z.; Chi, X.; Funabashi, M.; Baba, S.; Nonaka, K.; Pahari, P.; Unrine, J.; Jacobsen, J. M.; Elliott, G. I.; Rohr, J.; Van Lanen, S. G. Characterization of LipL as a non-heme, Fe(II)-dependent alpha-ketoglutarate:UMP dioxygenase that generates uridine-5′-aldehyde during A-90289 biosynthesis. *J. Biol. Chem.* **2011**, *286*, 7885–7892.

[66] Barnard-Britson, S.; Chi, X.; Nonaka, K.; Spork, A. P.; Tibrewal, N.; Goswami, A.; Pahari, P.; Ducho, C.; Rohr, J.; Van Lanen, S. G. Amalgamation of Nucleosides and Amino Acids in Antibiotic Biosynthesis: Discovery of an L-Threonine:Uridine-5′-Aldehyde Transaldolase. *J. Am. Chem. Soc.* **2012**, *134*, 18514–18517.

[67] Dutcher, J. D.; Vonsaltza, M. H.; Pansy, F. E. Septacidin, a New Antitumor and Antifungal Antibiotic Produced by Streptomyces Fibriatus. *Antimicrob. Agents Chemother.* **1963**, *161*, 83–88.

[68] Hayakawa, Y.; Nakagawa, M.; Kawai, H.; Tanabe, K.; Nakayama, H.; Shimazu, A.; Seto, H.; Otake, N. Studies on the Differentiation Inducers of Myeloid Leukemic-Cells .6. Spicamycin, a New Differentiation Inducer of Mouse Myeloid-Leukemia Cells (M1) and Human Promyelocytic Leukemia-Cells (Hl-60). *Agric. Biol. Chem.* **1985**, *49*, 2685–2691.

[69] Igarashi, Y.; Ootsu, K.; Onaka, H.; Fujita, T.; Uehara, Y.; Furumai, T. Anicemycin, a new inhibitor of anchorage-independent growth of tumor cells from *Streptomyces* sp TP-A0648. *J. Antibiot.* **2005**, *58*, 322–326.

[70] Harada, S.; Kishi, T. Isolation and Characterization of a New Nucleoside Antibiotic, Amipurimycin. *J. Antibiot.* **1977**, *30*, 11–16.

[71] Goto, T.; Toya, Y.; Ohgi, T.; Kondo, T. Structure of Amipurimycin, a Nucleoside Antibiotic Having a Novel Branched Sugar Moiety. *Tetrahedron Lett.* **1982**, *23*, 1271–1274.

[72] Seto, H.; Koyama, M.; Ogino, H.; Tsuruoka, T.; Inouye, S.; Otake, N. The Structures of Novel Nucleoside Antibiotics, Miharamycin A and Miharamycin B. *Tetrahedron Lett.* **1983**, *24*, 1805–1808.

[73] Cipolla, L.; Gabrielli, L.; Bini, D.; Russo, L.; Shaikh, N. Kdo: a critical monosaccharide for bacteria viability. *Nat. Prod. Rep.* **2010**, *27*, 1618–1629.

[74] Levin, D. H.; Racker, E. Condensation of arabinose 5-phosphate and phosphorylenol pyruvate by 2-keto-3-deoxy-8-phosphooctonic acid synthetase. *J. Biol. Chem.* **1959**, *234*, 2532–2539.

[75] Isono, K.; Crain, P. F.; Mccloskey, J. A. Isolation and Structure of Octosyl Acids - Anhydrooctose Uronic Acid Nucleosides. *J. Am. Chem. Soc.* **1975**, *97*, 943–945.

[76] Hori, M.; Eguchi, J.; Kakiki, K.; Misato, T. Studies on the mode of action of polyoxins. VI Effect of polyoxin B on chitin synthesis in polyoxin-sensitive and resistant strains of *Alternaria Kikuchiana*. *J. Antibiot.* **1974**, *27*, 260–266.

[77] Chen, W.; Huang, T.; He, X.; Meng, Q.; You, D.; Bai, L.; Li, J.; Wu, M.; Li, R.; Xie, Z.; Zhou, H.; Zhou, X.; Tan, H.; Deng, Z. Characterization of the polyoxin biosynthetic gene cluster from *Streptomyces cacaoi* and engineered production of polyoxin H. *J. Biol. Chem.* **2009**, *284*, 10627–10638.

[78] O'Connor, S.; Lam, L. K.; Jones, N. D.; Chaney, M. O. Apramycin, a unique aminocyclitol antibiotic. *J. Org. Chem.* **1976**, *41*, 2087–2092.

[79] Rinehart, K. L. Biosynthesis and Mutasynthesis of Aminocyclitol Antibiotics. *Abstracts of Papers of the American Chemical Society* **1979**, 12.

[80] Kudo, F.; Eguchi, T. Biosynthetic genes for aminoglycoside antibiotics. *J. Antibiot.* **2009**, *62*, 471–481.

[81] Komaki, E.; Ohta, Y.; Tsukada, Y. Purification and characterization of *N*-acetylneuraminate synthase from *Escherichia coli* K1-M12. *Biosci., Biotechnol., Biochem.* **1997**, *61*, 2046–2050.

[82] Comb, D. G.; Roseman, S. The sialic acids. I. The structure and enzymatic synthesis of N-acetylneuraminic acid. *J. Biol. Chem.* **1960**, *235*, 2529–37.

[83] Takahashi, S.; Nakagawa, F.; Sato, S. Griseolic acids B and C, an inhibitor of cyclic adenosine 3′,5′-monophosphate phosphodiesterase. *J. Antibiot.* **1988**, *41*, 705–706.

[84] Miyakoshi, S.; Haruyama, H.; Shioiri, T.; Takahashi, S.; Torikata, A.; Yamazaki, M. Biosynthesis of griseolic acids: incorporation of ^{13}C-labeled compounds into griseolic acid A. *J. Antibiot.* **1992**, *45*, 394–399.

[85] Harada, S.; Kishi, T. Isolation and Characterization of Mildiomycin, a New Nucleoside Antibiotic. *J. Antibiot.* **1978**, *31*, 519–524.

[86] Feduchi, E.; Cosin, M.; Carrasco, L. Mildiomycin - a Nucleoside Antibiotic That Inhibits Protein-Synthesis. *J. Antibiot.* **1985**, *38*, 415–419.

[87] Otake, N.; Takeuchi, S.; Endo, T.; Yonehara, H. Structure of Blasticidin S. *Tetrahedron Lett.* **1965**, 1411–1419.

[88] Wu, J.; Li, L.; Deng, Z.; Zabriskie, T. M.; He, X. Analysis of the Mildiomycin Biosynthesis Gene Cluster in *Streptoverticillum remofaciens* ZJU5119 and Characterization of MilC, a Hydroxymethyl cytosyl-glucuronic Acid Synthase. *ChemBioChem* **2012**, *13*, 1613–1621.

[89] Li, L.; Xu, Z.; Xu, X.; Wu, J.; Zhang, Y.; He, X.; Zabriskie, T. M.; Deng, Z. The Mildiomycin Biosynthesis: Initial Steps for Sequential Generation of 5-Hydroxymethylcytidine 5′-Monophosphate and 5-Hydroxymethylcytosine in *Streptoverticillium rimofaciens* ZJU5119. *ChemBioChem* **2008**, *9*, 1286–1294.

[90] Hamil, R. L.; Hoehn, M. M. A9145, a new adenine-containing antifungal antibiotic. I. Discovery and isolation. *J. Antibiot.* **1973**, *26*, 463–465.

[91] Berry, D. R.; Abbott, B. J. Incorporation of ^{14}C-labeled compounds into sinefungin (A9145), a nucleoside antifungal antibiotic. *J. Antibiot.* **1978**, *31*, 185–191.

[92] Takatsuki, A.; Arima, K.; Tamura, G. Tunicamycin, a New Antibiotic .1. Isolation and Characterization of Tunicamycin. *J. Antibiot.* **1971**, *24*, 215–223.

[93] Hamill, R. L. Process for preparing tunicamycin. US Patent US4237225, 1980.

[94] Takatsuki, A.; Kawamura, K.; Okina, M.; Kodama, Y.; Ito, T.; Tamura, G. Structural Elucidation of Tunicamycin .2. Structure of Tunicamycin. *Agric. Biol. Chem.* **1977**, *41*, 2307–2309.

[95] Tsvetanova, B. C.; Kiemle, D. J.; Price, N. P. Biosynthesis of tunicamycin and metabolic origin of the 11-carbon dialdose sugar, tunicamine. *J. Biol. Chem.* **2002**, *277*, 35289–35296.

[96] Wyszynski, F. J.; Hesketh, A. R.; Bibb, M. J.; Davis, B. G. Dissecting tunicamycin biosynthesis by genome mining: cloning and heterologous expression of a minimal gene cluster. *Chem. Sci.* **2010**, *1*, 581–589.

[97] Chen, W.; Qu, D.; Zhai, L.; Tao, M.; Wang, Y.; Lin, S.; Price, N. P.; Deng, Z. Characterization of the tunicamycin gene cluster unveiling unique steps involved in its biosynthesis. *Protein Cell* **2010**, *1*, 1093–1105.

[98] Wyszynski, F. J.; Lee, S. S.; Yabe, T.; Wang, H.; Gomez-Escribano, J. P.; Bibb, M. J.; Lee, S. J.; Davies, G. J.; Davis, B. G. Biosynthesis of the tunicamycin antibiotics proceeds via unique exo-glycal intermediates. *Nat. Chem.* **2012**, *4*, 539–546.

[99] Arai, M.; Haneishi, T.; Kitahara, N.; Enokita, R.; Kawakubo, K. Herbicidins A and B, two new antibiotics with herbicidal activity. I. Producing organism and biological activities. *J. Antibiot.* **1976**, *29*, 863–869.

[100] Chai, X.; Youn, U. J.; Sun, D.; Dai, J.; Williams, P.; Kondratyuk, T. P.; Borris, R. P.; Davies, J.; Villanueva, I. G.; Pezzuto, J. M.; Chang, L. C. Herbicidin Congeners, Undecose Nucleosides from an Organic Extract of *Streptomyces* sp. L-9-10. *J. Nat. Prod.* **2014**, *77*, 227–233.

[101] Yoshikawa, H.; Takiguchi, Y.; Terao, M. Terminal steps in the biosynthesis of herbicidins, nucleoside antiobiotics. *J. Antibiot.* **1983**, *36*, 30–35.

[102] Uchida, K.; Ichikawa, T.; Shimauch.Y,; Ishikura, T.; Ozaki, A. Hikizimycin, a New Antibiotic. *J. Antibiot.* **1971**, *24*, 259–262.

[103] Uchida, K. Structure of Hikizimycin, Nucleoside Antibiotic. *Agric. Biol. Chem.* **1976**, *40*, 395–404.

[104] Ennifar, S.; Das, B. C.; Nash, S. M.; Nagarajan, R. Structural Identity of Anthelmycin with Hikizimycin. *J. Chem. Soc., Chem. Commun.* **1977**, 41–42.

[105] Walsh, C. T.; Fischbach, M. A. Natural Products Version 2.0: Connecting Genes to Molecules. *J. Am. Chem. Soc.* **2010**, *132*, 2469–2493.

[106] Facchini, P. J. ALKALOID BIOSYNTHESIS IN PLANTS: Biochemistry, Cell Biology, Molecular Regulation, and Metabolic Engineering Applications. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2001**, *52*, 29–66.

[107] Lee, S. K.; Chou, H.; Ham, T. S.; Lee, T. S.; Keasling, J. D. Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Curr. Opin. Biotechnol.* **2008**, *19*, 556 – 563.

[108] Paddon, C. J. et al. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **2013**, *496*, 528–532.

[109] Lewis, C.; Clapp, H. W.; Grady, J. E. In vitro and in vivo evaluation of lincomycin, a new antibiotic. *Antimicrob. Agents Chemother.* **1963**, 570–582.

[110] Reusser, F. Effect of Lincomycin and Clindamycin on Peptide Chain Initiation. *Antimicrob. Agents Chemother.* **1975**, *7*, 32–37.

[111] Birkenmeyer, R. D.; Kagan, F. Lincomycin. XI. Synthesis and structure of clindamycin, a potent antibacterial agent. *J. Med. Chem.* **1970**, *13*, 616–619.

[112] Birkenmeyer, R. D.; Kroll, S. J.; Lewis, C.; Stern, K. F.; Zurenko, G. E. Synthesis and antimicrobial activity of clindamycin analogs: pirlimycin, a potent antibacterial agent. *J. Med. Chem.* **1984**, *27*, 216–223.

[113] Zhang, H.-Z.; Schmidt, H.; Piepersberg, W. Molecular cloning and characterization of two lincomycin-resistance genes, ImrA and ImrB, from *Streptomyces lincolnensls* 78-11. *Mol. Microbiol.* **1992**, *6*, 2147–2157.

[114] Brisson-Noël, A.; Delrieu, P.; Samain, D.; Courvalin, P. Inactivation of lincosaminide antibiotics in *Staphylococcus*. Identification of lincosaminide *O*-nucleotidyltransferases and comparison of the corresponding resistance genes. *J. Biol. Chem.* **1988**, *263*, 15880–15887.

[115] Peschke, U.; Schmidt, H.; Zhang, H.-Z.; Piepersberg, W. Molecular characterization of the lincomycin-production gene cluster of *Streptomyces lincolnensis* 78-11. *Mol. Microbiol.* **1995**, *16*, 1137–1156.

[116] Brahme, N. M.; Gonzalez, J. E.; Mizsak, S.; Rolls, J. R.; Hessler, E. J.; Hurley, L. H. Biosynthesis of the Lincomycins .2. Studies Using Stable Isotopes on the Biosynthesis of Methylthiolincosaminide Moiety of Lincomycin A. *J. Am. Chem. Soc.* **1984**, *106*, 7878–7883.

[117] Koběrská, M.; Kopecký, J.; Olšovská, J.; Jelínková, M.; Ulanova, D.; Man, P.; Flieger, M.; Janata, J. Sequence analysis and heterologous expression of the lincomycin biosynthetic cluster of the type strain *Streptomyces lincolnensis* ATCC 25466. *Folia Microbiol.* **2008**, *53*, 395–401.

[118] Čermák, L.; Novotná, J.; Ságová-Marečková, M.; Kopecký, J.; Najmanová, L.; Janata, J. Hybridization analysis and mapping of the celesticetin gene cluster revealed genes shared with lincomycin biosynthesis. *Folia Microbiol.* **2007**, *52*, 457–462.

[119] Cox, D. J.; Smith, M. D.; Fairbanks, A. J. Glycosylation Catalyzed by a Chiral Brønsted Acid. *Org. Lett.* **2010**, *12*, 1452–455.

[120] Czernecki, S.; Horns, S.; Valery, J.-M. Chain-Extension of Carbohydrates. 5. Synthesis of the *C*-Glycosyl Amino Acid Moiety of Miharamycins Involving Stereocontrolled Ethynylation of Methyl 2,3,4-Tri-*O*-benzyl-α-D-*gluco*-hexodialdo-1,5-pyranoside. *J. Org. Chem.* **1995**, *60*, 650–655.

[121] Yin, Z.-J.; Wang, B.; Li, Y.-B.; Meng, X.-B.; Li, Z.-J. Highly Efficient and Mild Method for Regioselective De-*O*-benzylation of Saccharides by $Co_2(CO)_8$-Et_3SiH–CO Reagent System. *Org. Lett.* **2010**, *12*, 536–539.

[122] Hoye, T. R.; Jeffrey, C. S.; Shao, F. Mosher ester analysis for the determination of absolute configuration of stereogenic (chiral) carbinol carbons. *Nat. Protoc.* **2007**, *2*, 2451–2458.

[123] Brahme, N. M.; Gonzalez, J. E.; Rolls, J. P.; Hessler, E. J.; Mizsak, S.; Hurley, L. H. Biosynthesis of the lincomycins. 1. Studies using stable isotopes on the biosynthesis of the propyl- and ethyl-L-hygric acid moieties of lincomycins A and B. *J. Am. Chem. Soc.* **1984**, *106*, 7873–7878.

[124] Neusser, D.; Schmidt, H.; Spížek, J.; Novotná, J.; Peschke, U.; Kaschabeck, S.; Tichy, P.; Piepersberg, W. The genes *lmbB1* and *lmbB2* of *Streptomyces lincolnensis* encode enzymes involved in the conversion of L-tyrosine to propylproline during the biosynthesis of the antibiotic lincomycin A. *Arch. Microbiol.* **1998**, *169*, 322–332.

[125] Colabroy, K. L.; Hackett, W. T.; Markham, A. J.; Rosenberg, J.; Cohen, D. E.; Jacobson, A. Biochemical characterization of L-DOPA 2,3-dioxygenase, a single-domain type I extradiol dioxygenase from lincomycin biosynthesis. *Arch. Biochem. Biophys.* **2008**, *479*, 131–138.

[126] Hu, Y.; Phelan, V.; Ntai, I.; Farnet, C. M.; Zazopoulos, E.; Bachmann, B. O. Benzodiazepine Biosynthesis in *Streptomyces refuineus*. *Chem. Biol.* **2007**, *14*, 691–701.

[127] Li, W.; Chou, S.; Khullar, A.; Gerratana, B. Cloning and Characterization of the Biosynthetic Gene Cluster for Tomaymycin, an SJG-136 Monomeric Analog. *Applied and Environmental Microbiology* **2009**, *75*, 2958–2963.

[128] Li, W.; Khullar, A.; Chou, S.; Sacramo, A.; Gerratana, B. Biosynthesis of Sibiromycin, a Potent Antitumor Antibiotic. *Appl. Environ. Microbiol.* **2009**, *75*, 2869–2878.

[129] Höfer, I.; Crüsemann, M.; Radzom, M.; Geers, B.; Flachshaar, D.; Cai, X.; Zeeck, A.; Piel, J. Insights into the Biosynthesis of Hormaomycin, An Exceptionally Complex Bacterial Signaling Metabolite. *Chem. Biol.* **2011**, *18*, 381–391.

[130] Kuo, M. S.; Yurek, D. A.; Coats, J. H.; Chung, S. T.; Li, G. P. Isolation and identification of 3-propylidene-Δ^1-pyrroline-5-carboxylic acid, a biosynthetic precursor of lincomycin. *J. Antibiot.* **1992**, *45*, 1773–1777.

[131] Kneidinger, B.; Graninger, M.; Puchberger, M.; Kosma, P.; Messner, P. Biosynthesis of Nucleotide-activatedd-*glycero*-D-*manno*-Heptose. *J. Biol. Chem.* **2001**, *276*, 20935–20944.

[132] Kneidinger, B.; Marolda, C.; Graninger, M.; Zamyatina, A.; McArthur, F.; Kosma, P.; Valvano, M. A.; Messner, P. Biosynthesis Pathway of ADP-L-*glycero*-β-D-*manno*-Heptose in *Escherichia coli*. *J. Bacteriol.* **2002**, *184*, 363–369.

[133] Newton, G. L.; Buchmeier, N.; Fahey, R. C. Biosynthesis and Functions of Mycothiol, the Unique Protective Thiol of Actinobacteria. *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 471–494.

[134] Newton, G. L.; Leung, S. S.; Wakabayashi, J. I.; Rawat, M.; Fahey, R. C. The DinB Superfamily Includes Novel Mycothiol, Bacillithiol, and Glutathione *S*-Transferases. *Biochemistry* **2011**, *50*, 10751–10760.

[135] Spížek, J.; Řezanka, T. Lincomycin, cultivation of producing strains and biosynthesis. *Appl. Microbiol. Biotechnol.* **2004**, *63*, 510–519.

[136] Fischbach, M. A.; Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev.* **2006**, *106*, 3468–3496.

[137] Kadlčík, S.; Kučera, T.; Chalupská, D.; Gažák, R.; Koběrská, M.; Ulanová, D.; Kopecký, J.; Kutejová, E.; Najmanová, L.; Janata, J. Adaptation of an L-Proline Adenylation Domain to Use 4-Propyl-L-Proline in the Evolution of Lincosamide Biosynthesis. *PLoS One* **2013**, *8*, e84902.

[138] Samland, A. K.; Rale, M.; Sprenger, G. A.; Fessner, W.-D. The Transaldolase Family: New Synthetic Opportunities from an Ancient Enzyme Scaffold. *ChemBioChem* **2011**, *12*, 1454–1474.

[139] Sasaki, E.; Lin, C.-I.; Lin, K.-Y.; Liu, H.-w. Construction of the Octose 8-Phosphate Intermediate in Lincomycin A Biosynthesis: Characterization of the Reactions Catalyzed by LmbR and LmbN. *J. Am. Chem. Soc.* **2012**, *136*, 906–909.

[140] Fitz, W.; Schwark, J.-R.; Wong, C.-H. Aldotetroses and C(3)-Modified Aldohexoses as Substrates for *N*-Acetylneuraminic Acid Aldolase: A Model for the Explanation of the Normal and the Inversed Stereoselectivity. *J. Org. Chem.* **1995**, *60*, 3663–3670.

[141] Barton, W. A.; Lesniak, J.; Biggins, J. B.; Jeffrey, P. D.; Jiang, J.; Rajashankar, K.; Thorson, J. S.; Nikolov, D. B. Structure, mechanism and engineering of a nucleotidylyltransferase as a first step toward glycorandomization. *Nat. Struct. Mol. Biol.* **2001**, *8*, 545–551.

[142] Sambrook, J.; Russell, D. *Molecular Cloning: A Laboratory Manual*, 3rd ed.; Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001.

[143] Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254.

[144] Peng, Z.-H.; Sharma, V.; Singleton, S. F.; Gershon, P. D. Synthesis and Application of a Chain-Terminating Dinucleotide mRNA Cap Analog. *Org. Lett.* **2002**, *4*, 161–164.

[145] Samuel, G.; Reeves, P. Biosynthesis of *O*-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and *O*-antigen assembly. *Carbohydr. Res.* **2003**, *338*, 2503–2519.

[146] Nedal, A.; Sletta, H.; Brautaset, T.; Borgos, S. E. F.; Sekurova, O. N.; Ellingsen, T. E.; Zotchev, S. B. Analysis of the Mycosamine Biosynthesis and Attachment Genes in the Nystatin Biosynthetic Gene Cluster of *Streptomyces noursei* ATCC 11455. *Appl. Environ. Microbiol.* **2007**, *73*, 7400–7407.

[147] Kong, D.; Lee, M.-J.; Lin, S.; Kim, E.-S. Biosynthesis and pathway engineering of antifungal polyene macrolides in actinomycetes. *J. Ind. Microbiol. Biotechnol.* **2013**, *40*, 529–543.

[148] Palaniappan, N.; Ayers, S.; Gupta, S.; Habib, E.-S.; Reynolds, K. A. Production of Hygromycin A Analogs in *Streptomyces hygroscopicus* NRRL 2388 through Identification and Manipulation of the Biosynthetic Gene Cluster. *Chem. Biol.* **2006**, *13*, 753–764.

[149] Du, L.; Sánchez, C.; Chen, M.; Edwards, D. J.; Shen, B. The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. *Chem. Biol.* **2000**, *7*, 623–642.

[150] Elbein, A. D. Trehalose Phosphate Synthesis in *Streptomyces hygroscopicus*: Purification of Guanosine Diphosphate D-Glucose: D-Glucose-6-Phosphate 1-Glucosyl-Transferase. *J. Bacteriol.* **1968**, *96*, 1623–1631.

[151] Butty, F. D.; Aucoin, M.; Morrison, L.; Ho, N.; Shaw, G.; Creuzenet, C. Elucidating the Formation of 6-Deoxyheptose: Biochemical Characterization of the GDP-D-*glycero*-D-*manno*-heptose C6 Dehydratase, DmhA, and Its Associated C4 Reductase, DmhB. *Biochemistry* **2009**, *48*, 7764–7775.

[152] Matthew, M.; Gary, S. S.; Carole, C. Characterization of the dehydratase WcbK and the reductase WcaG involved in GDP-6-deoxy-manno-heptose biosynthesis in *Campylobacter jejuni*. *Biochem. J.* **2011**, *439*, 235–248.

[153] McCallum, M.; Shaw, S. D.; Shaw, G. S.; Creuzenet, C. Complete 6-Deoxy-D-*altro*-heptose Biosynthesis Pathway from *Campylobacter jejuni*: MORE COMPLEX THAN ANTICIPATED. *J. Biol. Chem.* **2012**, *287*, 29776–29788.

[154] Kneidinger, B.; Graninger, M.; Adam, G.; Puchberger, M.; Kosma, P.; Zayni, S.; Messner, P. Identification of Two GDP-6-deoxy-D-*lyxo*-4-hexulose Reductases Synthesizing GDP-D-rhamnose in *Aneurinibacillus thermoaerophilus* L420-91T. *J. Biol. Chem.* **2001**, *276*, 5577–5583.

[155] Mäki, M.; Järvinen, N.; Räbinä, J.; Maaheimo, H.; Mattila, P.; Renkonen, R. Cloning and functional expression of a novel GDP-6-deoxy-D-talose synthetase from *Actinobacillus actinomycetemcomitans*. *Glycobiology* **2003**, *13*, 295–303.

[156] Albermann, C.; Piepersberg, W. Expression and identification of the RfbE protein from *Vibrio cholerae* O1 and its use for the enzymatic synthesis of GDP-D-perosamine. *Glycobiology* **2001**, *11*, 655–661.

[157] Rosano, C.; Bisso, A.; Izzo, G.; Tonetti, M.; Sturla, L.; Flora, A. D.; Bolognesi, M. Probing the catalytic mechanism of GDP-4-keto-6-deoxy-D-mannose epimerase/reductase by kinetic and crystallographic characterization of site-specific mutants. *J. Mol. Biol.* **2000**, *303*, 77–91.

[158] Alam, J.; Beyer, N.; Liu, H.-w. Biosynthesis of Colitose: Expression, Purification, and Mechanistic Characterization of GDP-4-keto-6-deoxy-D-mannose-3-Dehydrase (ColD) and GDP-L-colitose Synthase (ColC). *Biochemistry* **2004**, *43*, 16450–16460.

[159] Wolucka, B. A.; Van Montagu, M. GDP-mannose 3′,5′-epimerase forms GDP-L-gulose, a putative intermediate for the de novo biosynthesis of vitamin C in plants. *J. Biol. Chem.* **2003**, *278*, 47483–47490.

[160] Carmody, M.; Murphy, B.; Byrne, B.; Power, P.; Rai, D.; Rawlings, B.; Caffrey, P. Biosynthesis of Amphotericin Derivatives Lacking Exocyclic Carboxyl Groups. *J. Biol. Chem.* **2005**, *280*, 34420–34426.

[161] Hutchinson, E.; Murphy, B.; Dunne, T.; Breen, C.; Rawlings, B.; Caffrey, P. Redesign of Polyene Macrolide Glycosylation: Engineered Biosynthesis of 19-(*O*)-Perosaminyl-Amphoteronolide B. *Chem. Biol.* **2010**, *17*, 174–182.

[162] Silverman, R. B. *The organic chemistry of enzyme-catalyzed reactions*; Academic Press, 2002; Chapter 9, pp 387–340.

[163] Hallis, T. M.; Zhao, Z.; Liu, H.-w. New insights into the mechanism of CDP-D-tyvelose 2-epimerase: An enzyme-catalyzing epimerization at an unactivated stereocenter. *J. Am. Chem. Soc.* **2000**, *122*, 10493–10503.

[164] Zhao, Z.; Hong, L.; Liu, H.-w. Characterization of Protein Encoded by *spnR* from the Spinosyn Gene Cluster of *Saccharopolyspora spinosa*: Mechanistic Implications for Forosamine Biosynthesis. *J. Am. Chem. Soc.* **2005**, *127*, 7692–7693.

[165] Samuel, J.; Tanner, M. E. Mechanistic aspects of enzymatic carbohydrate epimerization. *Nat. Prod. Rep.* **2002**, *19*, 261–277.

[166] Ge, X.; Penney, L. C.; van de Rijn, I.; Tanner, M. E. Active site residues and mechanism of UDP-glucose dehydrogenase. *Eur. J. Biochem.* **2004**, *271*, 14–22.

[167] Daniellou, R.; Zheng, H.; Palmer, D. R. Kinetics of the reaction catalyzed by inositol dehydrogenase from *Bacillus subtilis* and inhibition by fluorinated substrate analogs. *Can. J. Chem.* **2006**, *84*, 522–527.

[168] Fontecave, M.; Ollagnier-de Choudens, S.; Mulliez, E. Biological radical sulfur insertion reactions. *Chem. Rev.* **2003**, *103*, 2149–2166.

[169] Song, H.; Her, A. S.; Raso, F.; Zhen, Z.; Huo, Y.; Liu, P. Cysteine Oxidation Reactions Catalyzed by a Mononuclear Non-heme Iron Enzyme (OvoA) in Ovothiol Biosynthesis. *Org. Lett.* **2014**, *16*, 2122–2125.

[170] Mashabela, G. T. M.; Seebeck, F. P. Substrate specificity of an oxygen dependent sulfoxide synthase in ovothiol biosynthesis. *Chem. Commun.* **2013**, *49*, 7714–7716.

[171] Tai, C.-H.; Cook, P. F. Pyridoxal 5′-phosphate-dependent α, β-elimination reactions: mechanism of *O*-acetylserine sulfhydrylase. *Acc. Chem. Res.* **2001**, *34*, 49–59.

[172] Mueller, E. G. Trafficking in persulfides: delivering sulfur in biosynthetic pathways. *Nat. Chem. Biol.* **2006**, *2*, 185–194.

[173] Mueller, E. G.; Palenchar, P. M.; Buck, C. J. The Role of the Cysteine Residues of ThiI in the Generation of 4-Thiouridine in tRNA. *J. Biol. Chem.* **2001**, *276*, 33588–33595.

[174] Loi, V. V.; Rossius, M.; Antelmann, H. Redox regulation by reversible protein *S*-thiolation in bacteria. *Front. Microbiol.* **2015**, *6*, *ASAP*.

[175] Vranish, J. N.; Russell, W. K.; Yu, L. E.; Cox, R. M.; Russell, D. H.; Barondeau, D. P. Fluorescent Probes for Tracking the Transfer of Iron-Sulfur Cluster and Other Metal Cofactors in Biosynthetic Reaction Pathways. *J. Am. Chem. Soc.* **2015**, *137*, 390–398.

[176] Scharf, D. H.; Remme, N.; Habel, A.; Chankhamjon, P.; Scherlach, K.; Heinekamp, T.; Hortschansky, P.; Brakhage, A. A.; Hertweck, C. A Dedicated Glutathione *S*-Transferase Mediates Carbon-Sulfur Bond Formation in Gliotoxin Biosynthesis. *J. Am. Chem. Soc.* **2011**, *133*, 12322–12325.

[177] Welch, T. R.; Williams, R. M. Epidithiodioxopiperazines. occurrence, synthesis and biogenesis. *Nat. Prod. Rep.* **2014**, *31*, 1376–1404.

[178] Zhang, C.; Griffith, B. R.; Fu, Q.; Albermann, C.; Fu, X.; Lee, I.-K.; Li, L.; Thorson, J. S. Exploiting the Reversibility of Natural Product Glycosyltransferase-Catalyzed Reactions. *Science* **2006**, *313*, 1291–1294.

[179] Schroeder, W.; Bannister, B.; Hoeksema, H. Lincomycin. III. Structure and stereochemistry of the carbohydrate moiety. *J. Am. Chem. Soc.* **1967**, *89*, 2448–2453.

[180] Knapp, S.; Kukkola, P. J. Stereocontrolled lincomycin synthesis. *J. Org. Chem.* **1990**, *55*, 1632–1636.

[181] Unson, M. D.; Newton, G. L.; Davis, C.; Fahey, R. C. An immunoassay for the detection and quantitative determination of mycothiol. *J. Immunol. Methods* **1998**, *214*, 29–39.

[182] Geoghegan, K. F.; Dixon, H. B.; Rosner, P. J.; Hoth, L. R.; Lanzetti, A. J.; Borzilleri, K. A.; Marr, E. S.; Pezzullo, L. H.; Martin, L. B.; LeMotte, P. K.; McColl, A. S.; Kamath, A. V.; Stroh, J. G. Spontaneous *α-N*-6-Phosphogluconoylation of a "His Tag" in *Escherichia coli*: The Cause of Extra Mass of 258 or 178 Da in Fusion Proteins. *Anal. Biochem.* **1999**, *267*, 169–184.

[183] Ulanova, D.; Novotná, J.; Smutná, Y.; Kameník, Z.; Gažák, R.; Šulc, M.; Sedmera, P.; Kadlčík, S.; Plháčková, K.; Janata, J. Mutasynthesis of lincomycin derivatives with activity against drug-resistant *staphylococci*. *Antimicrob. Agents Chemother.* **2010**, *54*, 927–930.

[184] Quadri, L. E.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **1998**, *37*, 1585–1595.

[185] Gehring, A. M.; Bradley, K. A.; Walsh, C. T. Enterobactin Biosynthesis in *Escherichia coli*: Isochorismate Lyase (EntB) Is a Bifunctional Enzyme That Is Phosphopantetheinylated by EntD and Then Acylated by EntE Using ATP and 2,3-Dihydroxybenzoate. *Biochemistry* **1997**, *36*, 8495–8503.

[186] Jia, L.; Tian, Y.; Tan, H. SanT, a bidomain protein, is essential for nikkomycin biosynthesis of *Streptomyces ansochromogenes*. *Biochem. Biophys. Res. Commun.* **2007**, *362*, 1031–1036.

[187] Zhao, Q.; Wang, M.; Xu, D.; Zhang, Q.; Liu, W. Metabolic coupling of two small-molecule thiols programs the biosynthesis of lincomycin A. *Nature* **2015**, *518*, 115–119.

Vita

Chia-I Lin was born on May 11, 1983 in Taitung, Taiwan. She graduated from National Taitung Girls' Senior High School in June, 2000 and attended the Department of Chemistry at National Taiwan University. After receiving her Bachelor's degree, Chia-I was admitted to the graduate program in the Department of Chemistry at National Taiwan University where she joined Professor Jim-Min Fang's research group. She is the first author of a paper regarding the mechanistic studies of pyreno[2,1-*b*]pyrroles as fluoride ion sensors and coauthored three papers and one patent describing the development of adenosine analogues for treating Huntington's Disease. After earning a master's degree in organic chemistry from National Taiwan University, Chia-I decided to pursue her doctoral studies in the University of Texas at Austin (Department of Chemistry, Organic Chemistry division) and joined the research group of Professor Hung-wen (Ben) Liu in 2008. She received the Graduate School Continuing Fellowship from the Department of Chemistry. Chia-I published three papers about her research on the biosynthetic studies of lincomycin A. Several additional research articles from her are currently in preparation.

Permanent email: captor.lin@gmail.com

This dissertation was typed by Chia-I Lin.